



UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Proposta de uma metodologia para a obtenção de vocabulários
de gestos intuitivos para a interação homem-robô

CLEBESON CANUTO DOS SANTOS

São Cristóvão

2016

CLEBESON CANUTO DOS SANTOS

**Proposta de uma metodologia para a obtenção de
vocabulários de gestos intuitivos para a interação
homem-robô**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências da Computação da Universidade Federal de Sergipe (PROCC-UFS), como requisito parcial para obtenção do título de Mestre em Ciência da Computação.

Área de concentração: Computação Inteligente

Orientador: Prof. Dr. Eduardo Oliveira Freire

São Cristóvão

2016

CLEBESON CANUTO DOS SANTOS

**Proposta de uma metodologia para a obtenção de
vocabulários de gestos intuitivos para a interação
homem-robô**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências da Computação da Universidade Federal de Sergipe (PROCC-UFS), como requisito parcial para obtenção do título de Mestre em Ciência da Computação.

Área de concentração: Computação Inteligente

BANCA EXAMINADORA

Prof. Dr. Eduardo Oliveira freire, Presidente
Universidade Federal de Sergipe - (PROCC/UFS)

Prof. Dr. Jugurta Rosa Montalvão Filho, Membro
Universidade Federal de Sergipe - (PROCC/UFS)

Prof. Dr. Lucas Molina, Membro
Universidade Federal de Sergipe - (PROEE/UFS)

Prof.^a Dr.^a Raquel Frizela Vassallo, Membro
Universidade Federal do Espírito Santo - (PPGEE/UFES)

**FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL
UNIVERSIDADE FEDERAL DE SERGIPE**

S237p Santos, Clebeson Canuto dos
Proposta de uma metodologia para a obtenção de vocabulários de gestos intuitivos para a interação homem-robô / Clebeson Canuto dos Santos ; orientador Eduardo Oliveira Freire. - São Cristóvão, 2016.
163 f. : il.

Dissertação (Mestrado em Ciência da Computação) - Universidade Federal de Sergipe, 2016.

1. Interação homem-máquina. 2. Linguística - processamento de dados. 3. Robótica. 4. Gestos - vocabulários. I. Freire, Eduardo Oliveira, orient. II. Título.

CDU 004.5:81'322

*Dedico este trabalho ao meu senhor Jesus Cristo que,
juntamente com o Pai Celestial e o Espírito Santo,
tem me dado toda a força e coragem necessária para
subir cada degrau e vencer cada barreira presente em
minha caminhada.*

Agradecimentos

A Deus por estar sempre ao meu lado, me erguendo e me dando forças para vencer as barreiras, mesmo quando elas parecem ser intransponíveis.

Aos meus pais, José Canuto dos Santos e Nailda Farias dos Santos, por me ensinarem a usar a coragem, a honestidade e a fé como ferramentas imprescindíveis para conseguir vencer os desafios.

A minha mãe e a minha noiva, Raily Vanessa Silva de Araújo, por estarem sempre ao meu lado nestes anos de estudos e trabalhos. Agradeço por toda a paciência que tiveram comigo, quando ficava até altas horas da madrugada com a luz acesa e fazendo barulho. Por serem compreensivas quando não participava de alguns eventos familiares porque estava estudando ou resolvendo exercícios. Por terem demonstrado paciência nos momentos de estresse e por me apoiarem em muitas outras situações em que precisei.

Ao meu orientador, o professor Dr. Eduardo Oliveira Freire, por toda a paciência demonstrada durante o processo de orientação. Por ter me mostrado a maneira correta de se fazer ciência. Por me ensinar que sempre temos algo a melhorar. Por me mostrar que para seguir carreira acadêmica não basta apenas ter um título, é necessário dedicar-se de maneira integral às pesquisas e se comprometer a passar da melhor maneira possível, todos os conhecimentos adquiridos para formar outros pesquisadores. Por dedicar vários finais de semanas e feriados corrigindo meus trabalhos e me guiando na pesquisa. Por ter me dado a oportunidade de participar do GPR-UFS, um grupo de pesquisa que considero ímpar. Agradeço ainda, por todos os conselhos e elogios dados, mas principalmente, por me repreender em determinados momentos e apontar meus erros de maneira humilde e educada, o que demonstra o seu comprometimento como orientador, professor e pesquisador.

Aos professores Dr. Elyson Ádan Nunes Carvalho, Dr. Lucas Molina e Dr. Jurgurta Rosa Montalvão Filho, por todos os conselhos e críticas construtivas dadas a cada apresentação dos resultados parciais. Tenham certeza de que o papel de vocês foi de fundamental importância para que os objetivos deste trabalho fossem alcançados.

Ao meu amigo Lucas Aragão de Carvalho, por todos os momentos de estudos, interação e debates.

Ao professor Dr. Almerindo Nascimento Rehem Neto, por me incentivar a ingressar na carreira acadêmica e por ter me apresentado ao professor Dr. Eduardo Oliveira Freire.

A todos os professores do Programa de Pós-Graduação em Ciência da Computação da UFS (PROCC/UFS) pela qualidade didática apresentada durante o primeiro ano de disciplinas.

A todos aqueles que fazem parte da coordenação do PROCC, em especial ao professor Dr. Rogério Patrício Chagas do Nascimento, por não medirem esforços para resolver problemas ou para responder meus e-mails, mesmo aos domingos e feriados.

A todos os integrantes da “família” GPR-UFS, por me ajudarem com os experimentos, por tirarem dúvidas de conteúdos técnicos e todo o apoio moral e afetivo dado durante a execução deste trabalho, meu muito obrigado.

Aos meus companheiros de mestrado, em especial os da linha de pesquisa em Computação Inteligente, por todos os momentos de estudos, debates e de descontração.

Por fim, a todos aqueles que, direta ou indiretamente, contribuíram para a realização deste trabalho.

"If language was given to men to conceal their thoughts, then gesture's purpose was to disclose them."

John Napier, Hands

Resumo

O desenvolvimento da robótica vem ganhando aceleração desde as últimas décadas, motivado principalmente pelo avanço da tecnologia, sobretudo dos computadores. No entanto, mesmo tendo tecnologia suficiente para criar robôs que possam participar do cotidiano das pessoas, a robótica ainda não se popularizou, haja vista que a maioria dos robôs adquiridos pelas pessoas ainda se enquadram na categoria de brinquedos, sistemas de monitoramento, dentre outros. A busca por esses tipos de robôs se deve ao fato de que seu repertório de tarefas é bem reduzido e predeterminado, o que acaba facilitando a interação entre usuários e robôs. Enquanto isso, robôs mais sofisticados, por possuírem um maior repertório de tarefas, acabam necessitando de mecanismos de interação mais complexos que, em sua maioria, só podem ser utilizados por pessoas especializadas. Em outras áreas, como a informática, a interface de comunicação foi de fundamental importância para a sua popularização. Dessa maneira, a criação de interfaces amigáveis de comunicação entre pessoas e robôs pode ser a chave para que a robótica também possa ser amplamente difundida na sociedade atual. No entanto, não é qualquer interface que pode oferecer uma comunicação fácil e eficiente. Para isso as mesmas devem ser o mais intuitivas possível, o que, segundo os estudos psicolinguísticos, pode ser alcançado por meio de gestos espontâneos. Logo, sabendo da dificuldade de se encontrar um procedimento que ofereça a possibilidade de se obter vocabulários de gestos intuitivos, esta dissertação de mestrado propõe uma metodologia que, baseada na psicolinguística e nos estudos sobre HCI (do inglês - *Human Computer Interaction*), oferece facilidade no processo de obtenção de vocabulários de gestos intuitivos a serem utilizados na interação entre pessoas e robôs. Desta maneira, após a aplicação de tal metodologia, pôde-se perceber que a mesma, apesar de ser nova, além de poder levar a resultados tão bons quanto uma outra metodologia já utilizada e aceita em HCI, ainda possui diferenciais, como a possibilidade de obter vocabulários mais complexos, que podem levar a vocabulários de gestos mais intuitivos e possivelmente mais robustos. Além disso, ao submeter os gestos obtidos a um reconhecedor, obteve-se uma taxa média de acertos de 77,5%, que mesmo não sendo alta, pode ser considerada boa, pois, uma vez que alguns gestos obtidos são realizados com os dois braços, a complexidade do reconhecimento é aumentada de maneira considerável. Assim, ao final desta dissertação, são apresentados vários trabalhos complementares a este, que devem ser realizados para que se possa avançar ainda mais na direção do desenvolvimento das interfaces intuitivas para a interação homem-robô.

Palavras-chave: Interação homem-robô, vocabulário de gestos, gestos intuitivos.

Abstract

Development in robotics has been accelerated in the last decades. Mainly due to the advancement in technology, especially computers. However, even having enough technology to create robots that can participate in the daily lives of people, robotics has not become popular, once that the most robots purchased by people still fall into the category of toys, monitoring systems, among others. The search for such robots is due to the fact that their repertoire of tasks is so much limited and predetermined, which ultimately facilitate the interaction between users and robots. Meanwhile, more sophisticated robots, in most cases, can only be used by specialized people, because they have a larger task repertoire, which needs more complex interaction mechanisms. In others areas, such as computing, the communication interface was of fundamental importance for its popularization. This way, the construction of friendly communication interfaces between people and robots can be the key to robotics can be widespread in the actual society. However, not every interface can provide an easy and efficient communication. An effective interface should be as intuitive as possible, what, according the psycholinguistics studies, can be achieved through the use of spontaneous gestures. Therefore, knowing the difficult to find a procedure to obtain intuitive gesture vocabularies, this master thesis proposes a methodology that, based on psycholinguistics and HCI (Human-Computer Interaction) studies, is suitable to obtain intuitive gesture vocabularies to be used in HRI (Human-Robot Interaction). Therefore, after the application of this methodology, it was possible to notice that it was able to lead to results as good as those obtained by another methodology which is already used and accepted in HCI. Moreover, the proposed methodology has some distinct characteristics, such as the possibility to obtain more complex vocabularies, that can lead to more intuitive gesture vocabularies and which, in form, are more likely to be robust. In addition, by submitting the obtained gestures to a recognizer, an average hit rate of 77,5% was obtained, which, even though it is not so high, can be considered good enough, since some of the gestures are performed with both arms, increasing the complexity of the recognition task. Thus, at the end of this master thesis, some complementary works are proposed, which must be carried out in order to move further towards the development of intuitive interfaces for human-robot interaction.

Keywords: *human-robot interaction, gesture vocabulary, intuitive gestures.*

Lista de figuras

Figura 1 – Exemplo de um gesto icônico (PRESS, 2012)	35
Figura 2 – Exemplo de um gesto metafórico (MCNEILL, 1992)	35
Figura 3 – Exemplo de um gesto dêitico (MATUSZEK et al., 2014)	36
Figura 4 – Exemplo de um gesto de batida (PAUL; PINKY, 2013)	37
Figura 5 – “Ok”: Gesto emblemático muito utilizado a nível mundial	38
Figura 6 – Gestos apresentados por Waldherr et al. (1998)	39
Figura 7 – Gestos apresentados por Meena, Jokinen e Wilcock (2012)	40
Figura 8 – Gestos apresentados por Barattini, Morand e Robertson (2012)	41
Figura 9 – Dicionário de gestos táteis apresentados por Yohanan e MacLean (2011)	42
Figura 10 – Gestos apresentados por Pereira (2012)	43
Figura 11 – Gestos apresentados por Fujii, Lee e Okamoto (2014)	44
Figura 12 – Gestos apresentados por Yang, Park e Lee (2007)	45
Figura 13 – Gestos apresentados por Wachs (2006)	47
Figura 14 – Etapas de realização da metodologia descrita em Medeiros, Tavares e Fonseca (2015)	51
Figura 15 – Modelo genérico do ambiente obtido após a análise dos gestos presentes em cada um dos vídeos	58
Figura 16 – <i>Microsoft Kinect Sensor 360</i> (MICROSOFT, 2015)	63
Figura 17 – Dados dos voluntários que realizaram os experimentos. Legenda: N/D - Não Definido, N.R - Não Responderam.	65
Figura 18 – Ambiente utilizado nos experimentos realizados no laboratório do GPR- UFS. A) visão do aplicador do experimento. B) visão do voluntário. C) visão de um observador externo	66
Figura 19 – Vocabulário 1: Vocabulário obtido com a aplicação da metodologia 1	70
Figura 20 – Vocabulário 2: Vocabulário obtido com a aplicação da metodologia 2	70
Figura 21 – Exemplo do alinhamento de duas séries quaisquer A e B	77
Figura 22 – Representação de um esqueleto virtual (CORP, 2016)	81
Figura 23 – Ambiente utilizado	105
Figura 24 – Gesto de recuar	109
Figura 25 – Vocabulário obtido com a aplicação da metodologia proposta	112

Lista de abreviaturas e siglas

ABF	Abordagem Baseada em Frustração
ANN	Redes Neurais Artificiais - do inglês <i>Artificial Neural Network</i>
DDF	Distribuição de Densidade de Característica - do inglês <i>Density Distribution Feature</i>
DSTW	Distorção Dinâmica no Espaço-Tempo - do inglês <i>Dynamic Space-Time Warping</i>
DTW	Distorção Dinâmica no Tempo - do inglês <i>Dynamic Time Warping</i>
EM	Expectation Maximization
FAPITEC	Fundação de Apoio à Pesquisa e Inovação Tecnológica
fdp	função de densidade de probabilidade
FSM	Máquina de Estados Finitos - do inglês <i>Finite States Machine</i>
GIV	Grau de intuitividade do Vocabulário
GPRUFS	Grupo de Pesquisa em Robótica da Universidade Federal de Sergipe
GMV	Grau de Memorização do Vocabulário
GSV	Grau de Semântica do Vocabulário
HCI	Interação Homem Computador - do inglês <i>Human-Computer Interaction</i>
HMM	Modelos Ocultos de Markov - do inglês <i>Hidden Markov Model</i>
HRI	Interação Homem Robô - do inglês <i>Human-Robot Interaction</i>
IA	Inteligência Artificial
ICA	Análise em Componentes Independentes - do inglês <i>Independent Component Analysis</i>
IDE	Ambiente de Desenvolvimento Integrado - do inglês <i>Integrated Development Environment</i>

KDES	Descritor de Núcleo - do inglês <i>Kernel Descriptor</i>
MHI	Histórico do Movimento de Imagens - do inglês <i>Motion History Images</i>
MIT	<i>Massachusetts Institute of Technology</i>
NASA	<i>National Aeronautics and Space Administration</i>
PCA	Análise de Componentes Principais - do inglês <i>Principal Component Analysis</i>
PG	Peso Global
RGB	Imagem colorida - do inglês <i>Red, Green, Blue</i>
RGB-D	Imagem colorida acrescida de suas informações de profundidade - do inglês <i>Red, Green, Blue and Depth</i>
s	segundos
SE	Estado de Sergipe
SQL	Linguagem de Consulta Estruturada - do inglês <i>Structured Query Language</i>
TDG	Tabela de Descrição dos Gestos
TGR	Tabela dos Gestos Realizados
SRI	<i>Stanford Research Institute</i>
SVM	Máquina de Vetores de Suporte - do inglês <i>Support Vector Machine</i>
TO	Taxas de Ocorrência
TOG	Taxa de Ocorrência Geral
TOT	Taxa de Ocorrência por Tempo
TOV	Taxa de Ocorrência por Voluntário
TPO	Taxa de Primeira Ocorrência
UNIT	Universidade Tiradentes
WoZ	Mágico de Oz - do inglês <i>Wizard of Oz</i>

Lista de Algoritmos

Algoritmo 1 – Algoritmo básico de um DTW	78
Algoritmo 2 – Algoritmo básico de um KNN	79

Lista de tabelas

Tabela 1 – Tabela do gestos estudados em Burke et al. (2004) (tradução própria) .	43
Tabela 2 – Quantidade de gestos para cada função usados na classificação.	82
Tabela 3 – Trabalhos que objetivam desenvolver reconhecedores de gestos	83
Tabela 4 – Funções escolhidas e os trabalhos onde elas aparecem.	104
Tabela 5 – Tabela de Descrição dos Gestos (TDG): Tabela onde todos os gestos encontrados com a aplicação dos experimentos são identificados e descritos	113
Tabela 6 – Tabela de Gestos Realizados (TGR): Todos os gestos realizados por cada um dos voluntários para cada tarefa.	118
Tabela 7 – Taxas de Ocorrência Geral (TOG) e pesos obtidos para cada gesto em cada função. As linhas estão ordenadas respectivamente pela função e pelo peso	139
Tabela 8 – Taxas de Primeira Ocorrência (TPO) e pesos obtidos para cada gesto em cada função. As linhas estão ordenadas respectivamente pela função e pelo peso	143
Tabela 9 – Taxas de Ocorrência por Voluntário (TOV) e pesos obtidos para cada gesto em cada função. As linhas estão ordenadas respectivamente pela função e pelo peso	147
Tabela 10 – Taxas de Ocorrência por Tempo (TOT) e pesos obtidos para cada gesto em cada função. As linhas estão ordenadas respectivamente pela função e pelo peso	152
Tabela 11 – Pesos de cada gesto em cada função. As linhas estão ordenados respectivamente pela função e pelo Peso Geral (PG)	157

Sumário

1	Introdução	20
1.1	Motivação	20
1.2	Definição do Problema	24
1.3	Hipótese Proposta	27
1.4	Objetivos	27
1.5	Metodologia	28
1.6	Recursos	30
1.7	Contribuições Esperadas	31
1.8	Estrutura da Dissertação	31
2	Estudo dos Gestos e sua Intuitividade	33
2.1	Estudo dos gestos segundo McNeill (1992)	33
2.1.1	Sobre os Gestos	33
2.1.2	Tipos de Gestos	34
2.1.2.1	Ícônicos	34
2.1.2.2	Metafóricos	35
2.1.2.3	Dêiticos (Apontar)	36
2.1.2.4	Batidas	36
2.1.2.5	Os gestos emblemáticos	37
2.2	Intuitividade dos Gestos	38
3	HRI baseada em gestos	39
3.1	Metodologias para obtenção de vocabulários de gestos	46
3.1.1	Metodologia proposta por Wachs	46
3.1.2	A metodologia de Nielsen	47
3.1.3	Outras Metodologias	49
3.2	Considerações	51
4	Metodologia Proposta	52
4.1	Escolha das funções	52
4.2	Captura dos dados gestuais	52

4.3	Análise dos gestos capturados	55
4.3.1	Análise dos gestos presentes em cada um dos vídeos . . .	55
4.3.2	Cálculos das taxas de ocorrência	56
4.4	Escolha do vocabulário	60
5	Aplicação da Metodologia Proposta	62
5.1	Etapa 1. Definição das Funções	62
5.2	Etapa 2. Captura dos dados gestuais	62
5.2.1	Definindo os parâmetros iniciais	62
5.2.2	Escolha do sensor	63
5.2.3	Os Voluntários	64
5.2.4	O Ambiente Utilizado	65
5.2.5	Os Experimentos	66
5.3	Etapa 3. Análise dos gestos	69
5.4	Etapa 4. Escolha do vocabulário	70
5.5	Aplicação da metodologia de Nielsen (metodologia 2)	71
5.6	Análise comparativa entre as metodologias 1 e 2	71
5.7	Classificação dos gestos dos vocabulários segundo a psicolinguística	73
5.7.1	Classificação dos gestos do vocabulário 1	73
5.7.2	Classificação dos gestos do vocabulário 2	74
5.7.3	Considerações sobre a classificação dos gestos	74
6	Reconhecimento dos Gestos Obtidos no Vocabulário	76
6.1	Conhecendo o DTW	76
6.2	Conhecendo o KNN	79
6.3	Reconhecimento dos gestos	80
6.3.1	Treinamento, testes e resultados	81
7	Conclusões e Trabalhos Futuros	85
7.1	Conclusões sobre o trabalho	85
7.2	Trabalhos Futuros	86
7.2.1	Ambiente de realização dos experimentos	86
7.2.2	Quantidade e perfil dos voluntários	87

7.2.3	Validação do vocabulário	88
7.2.3.1	Avaliação do grau de intuitividade do vocabulário de gestos	88
7.2.3.2	Experimentos práticos com o vocabulário obtido	90
7.2.4	Compartilhamento dos resultados	91
7.2.5	Integração entre fala e gestos	91
7.2.6	Inferência de contextos em ambientes interacionais	92
7.2.7	Melhoramento do modelo de reconhecimento dos gestos	92
7.2.8	Correlação entre trajetória, velocidade, número de re- petições e características únicas de um gesto	93
7.2.9	Interface de interação baseada em gestos universais	93
	Referências	95
	Apêndice A–Descrição detalhada da aplicação da metodologia	104
A.1	Etapa 1: Escolha das funções	104
A.2	Etapa 2: Captura dos dados Gestuais	105
A.2.1	Logística utilizada no ambiente	105
A.2.2	Etapas de Realização dos Experimentos	105
A.2.2.1	Apresentação	105
A.2.2.2	Orientações	106
A.2.2.3	Calibragem do <i>Kinect</i>	108
A.2.2.4	Demonstração	109
A.2.2.5	Início do experimento	110
A.2.3	Observação sobre as etapas realizadas	111
A.2.4	Etapa 3: Análise dos dados gestuais	111
A.2.5	Etapa 4: Escolha do vocabulário	111
	Apêndice B–Descrição dos gestos obtidos nos ex- perimentos	113
	Apêndice C–Gestos de cada voluntário em cada tarefa.	118

Apêndice D–Tabelas com as taxas de ocorrências dos gestos	139
Apêndice E–Pesos dos gestos em cada função . .	157
Apêndice F–Documentos utilizados nos experi- mentos	162

1 Introdução

1.1 Motivação

Existem relatos de que em torno do ano 2000 A.C. já se tentava construir ajudantes mecânicos para ajudar as pessoas nas tarefas diárias. Há indícios de que os egípcios criaram um cachorro mecânico e também que gregos relatavam a criação, por parte dos deuses, de mulheres de ouro, montadas em tripés com rodas, podendo assim andar entre os seres humanos e se comunicarem (PEREIRA, 2012). No ano de 1474, Leonardo Da Vinci criou um leão mecânico para agradar ao rei da França, e em 1495 ele desenhou um cavaleiro (a princípio para ser feito de metal) capaz de sentar, mexer os braços, cabeça, além do maxilar, sendo um primeiro vestígio da tentativa de criação de um humanoide. Segundo Chevallereau et al. (apud AKHTARUZZAMAN Md; SHAFIE A. A., 2010) em 1865, em uma obra de ficção de John Brainerd, foi criado um humanoide movido a vapor (*Steam Man* – em inglês), que ajudava os seres humanos carregando-os em carroças como se fora um cavalo ou um outro animal de carga. Essa obra inspirou o escritor Rank Reade Junior a publicar, em 1885, uma outra obra de ficção, na qual, o “*Steam Man*” não era movido à vapor, e sim à eletricidade e por isso foi chamado “*Electric Man*”.

Mesmo sabendo da importância de se ter uma “criatura” não humana capaz de realizar trabalhos braçais, e após todas as histórias e relatos de tentativas de criação de tal ser, antes do século 20 não se tinha um termo próprio para caracterizar esse tipo de dispositivo. Assim, em 1921 em uma peça de ficção intitulada “*Rossum’s Universal Robots*”, escrita pelo tcheco Karel Capek, surgiu o termo “robô” (CAPEK et al., 1920). O termo foi uma derivação da palavra “*robota*” que em tcheco significa “trabalho forçado” e foi difundido pelo mundo até os dias atuais.

Na década de 1960, a robótica ganhou mais espaço devido ao interesse no uso industrial (ENGELBERGER, 1995). Aumentou o interesse pela substituição do homem por máquinas, a fim de automatizar o processo de fabricação de diversos produtos, além da realização de tarefas perigosas (PEREIRA, 2012). Assim, em 1960, surgiu o *Unimate*, braço robótico implantado pela General Motors a fim de diminuir acidentes causados no processo de empilhamento de metais fundidos e no processo de soldagem das portas dos veículos, o qual foi considerado o primeiro robô a ser implantado efetivamente na indústria (MORAN, 2007). Ainda na década de 1960, mais especificamente no ano de 1966, nos laboratórios

da *Stanford Research Institute* (SRI), foi criado o primeiro robô a usar de maneira bem sucedida a Inteligência Artificial (IA). O Shakey ([BROOKS, 1991](#)), como assim foi chamado, era um robô móvel capaz de criar mapas de ambientes, abrir e fechar portas, ligar e desligar lâmpadas, além de executar muitas outras tarefas. Essa capacidade de tomada de decisão fornecida por meio da IA acabou abrindo caminho para os robôs autônomos que, a partir da década de 1970, tiveram um maior desenvolvimento.

No final da década de 1980, no ano de 1989, foi criado pelo MIT (*Massachusetts Institute of Technology*) o Genghis ([ANGLE, 1989](#)), robô com seis patas anatomicamente semelhante a uma barata. Este foi o primeiro robô a ter a capacidade de andar de maneira eficiente por diferentes tipos de superfícies, inspirando a criação de robôs como o BigDog ([RAIBERT et al., 2008](#)), capaz de se locomover em diversos tipos de ambientes através do movimento de pernas e não de rodas, como era comum observar na robótica móvel.

Na década de 1990, a criação de robôs cada vez mais inteligentes foi se acentuando. Dentre os muitos que surgiram, podem-se citar alguns mais importantes, como o Pioneer, um dos robôs móveis mais utilizados em pesquisas ([SIEGWART; NOURBAKHS; SCARAMUZZA, 2011](#)); o Sojourner, robô criado pela NASA e usado na missão “*Mars Pathfinder*”, que teve a tarefa de explorar a superfície marciana e analisar os compostos de alguns tipos de rochas ([VOLPE et al., 1996](#)); o Lego *Mindstorm* ([KLASSNER; ANDERSON, 2001](#)), *kit* robótico criado pela Lego com o intuito de facilitar o aprendizado da robótica, o qual era formado por conjunto de componentes pré-moldados, juntamente com *frameworks* que facilitavam a programação, deixando de lado a necessidade do usuário ser especialista em *hardware* e/ou *software* para construir seu próprio robô, permitindo assim que a robótica pudesse chegar de maneira mais fácil a ambientes acadêmicos e escolares; e já no final daquela década, em 1999, a Sony criou o Aibo ([TOKYO; NARA, 2004](#)), cão robô capaz de interagir e aprender com seu dono e que foi o primeiro robô a ser vendido em larga escala para ambientes domésticos, pois poderia ser usado como brinquedo.

No início dos anos 2000, tentando cumprir a promessa feita em 1986 de criar um robô que pudesse interagir através de características semelhantes às humanas, a Honda lançou o Asimo ([SAKAGAMI; WATANABE; AOYAMA, 2002](#)), robô humanoide capaz de realizar gestos manuais, andar, acenar com a cabeça, escutar, falar e muitas outras ações que se assemelham às de uma pessoa. Mesmo ainda não realizando movimentos de maneira tão perfeita como os de um ser humano, o surgimento do Asimo serve para mostrar que a robótica alcançou um nível de desenvolvimento capaz de poder ser incorporada à vida das

pessoas. Em 2002, a iRobot lança o Roomba ([JONES, 2006](#)), aspirador de pó autônomo capaz limpar pisos e carpetes de maneira muito eficaz. Alguns anos depois, após algumas atualizações, o Roomba tornou-se um dos robôs domésticos mais populares do mundo. Tal popularidade se deu devido à capacidade que o mesmo tem de limpar cômodos sem que seja necessária a supervisão de um ser humano para livrá-lo de obstáculos ou colocá-lo para recarregar as baterias, já que o mesmo possui bons mecanismos para evitar obstáculos, detectar sujeira e identificar e retornar para sua base de recarga assim que termina suas tarefas.

Além de todos os robôs supracitados, nas últimas décadas surgiram muitos outros robôs que tiveram uma importante relevância para o desenvolvimento da robótica, como o Robonaut ([BLUETHMANN et al., 2003](#)), humanoide criado em 2002 através de parceria da NASA com a General Motors, a fim de ser usado para auxiliar astronautas em missões espaciais, e o NAO ([KULK; WELSH, 2005](#)), humanoide programável e compacto criado em 2006 pela Aldebaran Robots, capaz de reproduzir movimentos humanos e muito utilizado em pesquisas científicas.

Surgiram também os robôs “*babysitters*” e os “*caregivers*” que são respectivamente robôs babás e cuidadores. Os *caregivers*, como o Giraffplus ([CORADESCHI et al., 2013](#)) e o RP-VITA ([KRISTOFFERSSON; CORADESCHI; LOUTFI, 2013](#)), são robôs teleoperados que se comunicam com o usuário através de uma tela tátil. A aplicação principal para esse tipo de robô é no auxílio ao cuidado de pessoas idosas, enfermas ou com algum tipo de deficiência. Certamente, uma tela tátil, por mais que seja adequada para utilizar um *tablet*, não é a forma como as pessoas costumam se comunicar. Dessa forma, a interação dos principais usuários com esses robôs fica bastante prejudicada, e frequentemente demanda a presença de um profissional especializado para operar o robô e orientar o usuário ([DAHL; BOULOS, 2013](#)). Ainda na categoria de *caregivers*, existem robôs como o Ri-Man ([ONISHI et al., 2007](#)) e o RIBA ([MUKAI et al., 2010](#)) que ajudam no intercâmbio de pessoas entre camas e cadeiras-de-rodas. Nesse contexto, as interfaces também são táteis, mas, já percebendo a importância de uma interface de comunicação intuitiva, existem pesquisas para utilizar a fala e/ou os gestos como interfaces de interação. Como exemplo de *babysitter* existe o T2-5 RUKURU ([TMSUK, 2009](#)), um robô criado pela empresa japonesa Tmsuk e vendido em supermercados, que permite identificar crianças através do reconhecimento de uma imagem *QR-Code* presente em crachás, conversar com as crianças (mesmo tendo um reduzido repertório de palavras), apresentar avisos pré-definidos pelos pais através de um projetor

presente em um de seus olhos, dentre outras tarefas. A principal desvantagem encontrada ao usar um destes robôs é que as tarefas a serem executadas devem ser pré-definidas pelos pais através de comandos criados pelo fabricante, o que reduz muito a sua versatilidade e aplicabilidade.

Dessa forma, pode-se perceber que o desenvolvimento da robótica vem ganhando aceleração desde as últimas décadas, motivada principalmente pelo avanço da tecnologia, sobretudo dos computadores. No entanto, mesmo tendo tecnologia suficiente para criar robôs que possam participar do cotidiano das pessoas, a robótica ainda não se popularizou, haja vista que a maioria dos robôs adquiridos pelas pessoas ainda se enquadram na categoria de brinquedos, sistemas de monitoramento, dentre outros. A busca por esses tipos de robôs se deve ao fato de que seu repertório de tarefas é bem reduzido e predeterminado, o que acaba facilitando a interação entre usuários e robôs. Enquanto isso, robôs mais sofisticados, por possuírem um maior repertório de tarefas, acabam necessitando de mecanismos de interação mais complexos que, em sua maioria, só podem ser utilizados por pessoas especializadas. Assim, tais robôs acabam sendo adquiridos apenas por indústrias e pesquisadores, já que a interação com os mesmos exige conhecimentos técnicos de programação e/ou robótica que a população em geral não possui.

Em outras áreas, a interface de comunicação foi de fundamental importância para a sua popularização. Tomando como exemplo os computadores, vê-se que em meados de 1970, mesmo tendo tecnologia de *hardware* e *software* suficiente para baratear custos, os computadores, em sua maioria, ainda eram utilizados apenas para fins científicos e corporativos, uma vez que a comunicação entre homem e máquina se dava de maneira muito complexa. Assim, com o surgimento dos sistemas operacionais gráficos e do *mouse*, a interação homem-computador passou de linhas de comando para o clique de um botão, fazendo com que pessoas menos instruídas tecnologicamente pudessem utilizar os computadores em seu cotidiano. Atualmente, com o surgimento das telas táteis, até mesmo crianças em idade pré-escolar podem interagir de maneira eficaz com um *tablet* ou *smartphone*. Desta maneira, o desenvolvimento de interfaces naturais e/ou intuitivas, influenciou diretamente na popularização dos computadores.

Em virtude de tudo que foi mencionado, assim como aconteceu com a informática, fica claro que a falta de uma interface amigável entre pessoas e robôs é um obstáculo que precisa ser removido para que a robótica possa ser amplamente difundida na sociedade atual.

1.2 Definição do Problema

A segunda lei de Asimov diz que um robô deve ser capaz de receber e realizar ordens humanas ([ASIMOV, 1950](#)). Mesmo em se tratando de uma lei proposta em uma obra de ficção, ela permite pensar na existência de um mecanismo/interface de comunicação que permita a interação entre o homem e o robô. Existem várias interfaces de comunicação homem-robô (HRI – do Inglês *Human-Robot Interaction*) sendo que as principais são: voz ([ROGALLA et al., 2002](#)), controles manuais ([GALLOWAY; RYU; AGRAWAL, 2008](#)), controles mentais ([NICOLELIS; CHAPIN, 2002](#)) e gestos ([LEE, 1996](#)).

Com as interfaces baseadas em controles manuais, a interação pode ser bem mais intuitiva do que com outras, pois, por exemplo, uma pessoa que controla um robô com um *joystick* sabe que o robô deverá seguir no mesmo sentido e orientação do posicionamento da alavanca ([LEE, 1996](#)). Entretanto, o uso de controles manuais, como o *joystick*, está limitado à teleoperação de um robô, não sendo adequado para que um operador dê ordens que devam ser executadas pelo robô de forma autônoma. Por outro lado, em interfaces como voz/fala e gestos, os comandos que deverão ser realizados para controlar o robô já não são tão intuitivos assim, e por isso necessitam de uma atenção especial na hora de sua concepção.

Segundo [Breazeal e Aryananda \(2002\)](#), além de ser uma interface intuitiva para comunicação entre homem e robô, a fala é muito eficiente no processo de aprendizagem de máquina. Isso porque a acústica oferece a possibilidade de reconhecer quem está falando, saber na íntegra o que se disse, e identificar a maneira como foi dito, possibilitando que o robô identifique com quem está interagindo e saber o que a pessoa deseja, através do que foi dito e da sua entonação da voz. Sendo assim, muitos trabalhos como [Aly e Tapus \(2012\)](#), [Buss et al. \(2011\)](#) abordam o reconhecimento de fala como interface de comunicação para interação homem-robô.

Dentre as interfaces previamente citadas, aquelas baseadas em gestos estão entre as que mais empregam esforço computacional, devido a que um gesto é resultado de um conjunto de poses dentro de um intervalo de tempo. Assim, para reconhecê-los, os sistemas reconhecedores recebem como entrada imagens, medidas de distância, informações de profundidade, dentre outros, sendo que em sua maioria são dados em grande quantidade. Por isso, muitas vezes os pesquisadores preferem optar, por exemplo, por comandos de voz que podem ser um pouco menos custosos do ponto de vista computacional.

[McNeill \(1992\)](#) explana que os gestos e a fala são co-expressivos, ou seja, ambos surgem de uma mesma fonte semântica compartilhada, mas não são redundantes, um complementa o outro. Pensando nisso, uma interface baseada em gestos e fala pode trazer mais informação para o robô e se tornar ainda mais intuitiva para o usuário do que aquela que se baseia apenas na fala. Mas existe um porém: o que pode ser intuitivo para alguém, pode não ser intuitivo para outra pessoa que não se comunica por meio da mesma linguagem a qual o robô foi programado para entender e interagir.

[Kendon \(2007\)](#) explana sobre a possível padronização dos gestos. Ele crê que existe um padrão para os gestos e que as pessoas, mesmo de diferentes culturas, compartilham esse conhecimento. Assim, no trabalho de [Goldin-Meadow et al. \(2008\)](#) eles concluíram, através de experiências envolvendo pessoas de diferentes países, culturas e línguas, que na ausência de fala, os gestos têm a mesma ordem de execução, o que indica a possibilidade de que uma interface intuitiva baseada apenas em gestos possa ser utilizada para comunicação com um robô em qualquer parte do mundo. Dessa maneira, a criação da interface de comunicação independe da pessoa que a utilizará. Outro aspecto a favor do uso dos gestos é que para [McNeill \(1992\)](#) eles possuem tanto significado quanto a fala. Por isso, é possível que interfaces baseadas apenas em gestos possam ser tão intuitivas quanto aquelas baseadas em voz.

[Wexelblat \(1998\)](#) afirma em seu trabalho que o estudo dos gestos para a comunicação homem-máquina estava um tanto estagnado, e que a comunidade científica deveria dar mais importância para essa área, pois o gesto é uma forte interface de comunicação. Wexelblat ainda diz que a falta de interesse pelo estudo dos gestos pode fazer com que essa área de pesquisa seja esquecida e torne a vida de seus pesquisadores mais difícil, pois, segundo ele, a cada novo trabalho envolvendo gestos, se empregam experimentos exaustivos devido à falta de bons trabalhos anteriores.

Após análise de diversos trabalhos sobre HRI, percebe-se que muitos deles mostram preocupação na utilização dos gestos como interface de comunicação, porém não trazem um estudo detalhado de como os gestos utilizados em seus experimentos foram escolhidos e/ou quão efetivos são esses gestos. Dessa maneira, os gestos utilizados ficam atrelados ao uso somente das pessoas que realizaram as pesquisas e/ou fizeram um determinado treinamento. Logo, na visão de Wexelblat, o uso de gestos nesse contexto deixa de ser vantajoso, pois, segundo ele, caso não haja preocupação com a naturalidade da interface de interação, seria mais atrativo utilizar controles manuais do que gestos ou fala, já que as

taxas de acerto na maioria das interfaces manuais chegam a 100%, além de serem bem menos custosas do ponto de vista do projetista e computacional.

Na tentativa de estudar os gestos táteis, [Yohanan e MacLean \(2011\)](#) se deparam com a dificuldade de encontrar na literatura alguns gestos desse tipo já predefinidos para que outras pesquisas pudessem utilizá-los. Sendo assim, criaram um dicionário com 30 gestos táteis no intuito de diminuir o esforço empregado em seus futuros trabalhos, ou evitar que esse esforço tivesse que ser repetido por outros pesquisadores. No entanto, mesmo se preocupando com o fornecimento do dicionário de gestos para futuras pesquisas, o trabalho não aborda qual a metodologia utilizada na obtenção de cada gesto que compõe o dicionário. Assim como este, muitos outros trabalhos, como ([URIBE; ALVES, 2011](#); [HASANUZZAMAN et al., 2005](#); [YANG; PARK; LEE, 2007](#); [WALDHERR et al., 1998](#)), mostram os gestos utilizados, mas não apresentam qual o embasamento que foi utilizado para a concepção dos mesmos, e outros, como ([BREMNER et al., 2009](#)), deixam claro que os gestos foram escolhidos de maneira empírica, o que gera desconfiança sobre sua real efetividade.

Desta maneira, a cada nova pesquisa se faz necessário buscar uma metodologia para a obtenção do vocabulário de gestos, sendo que o tempo destinado para isso poderia ser economizado, caso os pesquisadores se preocupassem mais em viabilizar o uso dos seus resultados por outros pesquisadores. Assim, é de fundamental importância que os trabalhos sobre HRI baseados em gestos sejam mais completos, no que diz respeito à descrição das tarefas realizadas pelo robô, aos gestos escolhidos para representar cada uma das tarefas e à metodologia empregada para a escolha de tais gestos. É importante ainda que a metodologia utilizada seja descrita de maneira a facilitar sua reprodução por outros pesquisadores, para que assim seja possível ter referenciais para comparação de resultados, o que pode acelerar o desenvolvimento dessa área de pesquisa.

Por esse motivo, [Stern, Wachs e Edan \(2006\)](#) resolveram buscar na área de HCI uma metodologia que pudesse ser utilizada na busca de vocabulários intuitivos. Esta foi uma decisão acertada, uma vez que, segundo ([POWERS, 2008](#)), como a área de HCI possui regras e conceitos bem definidos e aceitos pela comunidade acadêmica, ela pode oferecer soluções para alguns problemas ainda não resolvidos na área de HRI. Assim, a metodologia mais clara encontrada foi aquela descrita por [Nielsen et al. \(2004\)](#), que define uma sequência de passos a serem seguidos para se conseguir um vocabulário de gestos intuitivo e ergonômico. Porém, mesmo [Stern, Wachs e Edan \(2006\)](#) e [Stern et al. \(2008\)](#) considerando-a uma metodologia adequada para quem busca gestos intuitivos, o

embasamento científico no qual ela se sustenta para a determinação da intuitividade dos gestos é inconsistente, pois a maioria dos estudos e definições citados se baseiam quase que totalmente na engenharia de usabilidade ao invés da psicolinguística, sendo esta última mais indicada para esse tipo de estudo.

Diante disso, o problema a ser estudado nesta dissertação de mestrado consiste no desenvolvimento de uma metodologia que, à luz dos estudos já realizados a respeito de psicolinguística e baseando-se em uma metodologia já consolidada pela área de HCI, possa oferecer um caminho de fácil reprodução para a escolha de vocabulários de gestos intuitivos para interface homem-robô.

1.3 Hipótese Proposta

Como hipótese de trabalho, acredita-se que o uso dos conhecimentos oriundos do estudo de psicolinguística juntamente com os conhecimentos da área de HCI, permitirão o desenvolvimento de uma metodologia para se obter vocabulários de gestos intuitivos para interface homem-robô, que sejam significativamente diferentes entre si, facilitando assim a tarefa de reconhecimento dos mesmos e melhorando a taxa de sucesso do processo de interação entre usuários e robôs.

A solução para o problema apresentado e a confirmação da hipótese levantada devem resultar em uma contribuição para o desenvolvimento de trabalhos futuros na área de HRI baseada em gestos, além de permitir o desenvolvimento de uma interface homem-robô funcional e intuitiva, desenvolvida com o devido embasamento científico, e que pode vir a contribuir para a popularização da robótica na sociedade.

1.4 Objetivos

O objetivo geral desse trabalho é propor uma metodologia que, baseada na psicolinguística e nos estudos da área de HCI, ofereça facilidade na obtenção de vocabulários de gestos intuitivos que possam ser utilizados na interação entre homens e robôs.

Para atingir o objetivo geral, os seguintes objetivos específicos foram traçados:

1. Realizar uma ampla revisão bibliográfica visando conhecer o estado da arte na área de HRI.

2. Levantar e estudar os trabalhos da psicolinguística que falam dos gestos, para tentar encontrar uma relação científica entre a realização dos gestos e sua intuitividade.
3. Levantar e estudar os trabalhos sobre HCI que falam de gestos intuitivos, a fim de encontrar uma metodologia que possa servir de base para a proposta deste trabalho.
4. Conceber, por meio dos estudos realizados, uma metodologia de fácil reprodução que possa ser empregada na obtenção de vocabulários de gestos intuitivos.
5. Quantificar nos trabalhos levantados sobre HRI quais são as principais funções que os robôs executam, para que sejam utilizadas como base para a aplicação da metodologia proposta.
6. Conceber, baseando-se na metodologia proposta e nas funções levantadas, um conjunto de gestos de maneira que se tenha pelo menos um gesto para cada tarefa, e que sejam tão intuitivos quanto possível.
7. Escolher e implementar um sistema de reconhecimento de gestos baseado nos dados de esqueleto virtual, fornecido por um sensor RGB-D (imagem colorida mais profundidade), para realizar a identificação dos gestos concebidos.

1.5 Metodologia

A primeira etapa deste trabalho ocorreu a partir de estudos dirigidos que resultaram na definição do tema da pesquisa. Foi realizado um levantamento bibliográfico sobre os trabalhos desenvolvidos sobre o uso dos gestos em HRI e dos gestos em geral, a fim de contextualizar a problemática e de definir os objetivos da pesquisa.

Pensando nas indagações que o trabalho se propunha a responder, foi realizada uma pesquisa bibliográfica que serviu de referencial teórico para assim dar comprovações científicas sobre o que será abordado no decorrer do texto por meio das citações. Esta etapa prosseguiu paralelamente a todas às outras etapas da pesquisa até a conclusão desta dissertação. Assim foi possível mapear o máximo de trabalhos que surgiram nesta mesma linha de pesquisa, o que enriqueceu ainda mais o resultado final do presente trabalho.

Baseando-se nos estudos realizados e nos problemas encontrados nos trabalhos sobre HRI, foi concebida uma metodologia que possibilita a obtenção de gestos intuitivos que podem compor um vocabulário de gestos para interação homem-robô. Durante o desenvolvimento desta etapa, percebeu-se que na área de HCI já havia uma metodologia que se propunha a encontrar vocabulários, no entanto, ao analisá-la, percebeu-se que

além de utilizar a ergonomia como um dos fatores de escolha dos gestos, tal metodologia também restringe o vocabulário de maneira a permitir somente gestos estáticos e com relação de apenas um gesto para cada função, o que não maximiza a importância da intuitividade para a interface. Assim, a metodologia aqui proposta, além de se basear apenas na intuitividade, possui a característica de poder encontrar vocabulários de gestos onde uma função a ser realizada pelo robô pode ser representada por mais de um gesto.

Objetivando abordar com maior propriedade científica os experimentos realizados, os quais seguiram as etapas descritas na metodologia proposta, foram mapeadas as principais funções executadas pelos robôs mencionados nos trabalhos sobre HRI anteriormente analisados.

Os experimentos foram realizados por meio da captura de imagens de voluntários. Cada voluntário, antes de começar o experimento, assinou um termo de consentimento, no qual autorizava, ou não, que suas imagens aparecessem em eventuais publicações dos resultados deste trabalho. Sendo assim, todos os experimentos foram realizados com pessoas maiores de 18 anos. Para isso, toda essa etapa foi realizada em 5 fases, a saber:

1. Inicialmente, com o intuito de definir alguns padrões, como o tempo de realização de cada experimento e a melhor distância para captura dos dados, foram realizados experimentos piloto com 7 integrantes do Grupo de Pesquisa em Robótica da Universidade Federal de Sergipe (GPR-UFS).
2. Com autorização do coordenador do curso de Mecatrônica da Universidade Tiradentes - UNIT/SE, foram realizados mais de 20 experimentos em uma única ocasião.
3. Durante a Semana de Engenharia de 2015, promovida pela Faculdade Pio Décimo, foram realizados mais de 20 experimentos com alunos de diversos cursos daquela instituição.
4. Em nova oportunidade na Universidade Tiradentes, foram realizados quase 30 experimentos com alunos de diversos cursos.
5. Por fim, durante a Feira de Ciências promovida pela FAPITEC/SE na Semana Nacional de Ciência e Tecnologia de 2015, realizada Universidade Federal de Sergipe, foram realizados mais 10 experimentos, os quais, somados com os experimentos anteriores, totalizaram 84.

Após a realização dos experimentos e ainda seguindo a metodologia proposta, os dados foram analisados a fim de catalogar os gestos realizados e tornar factível a obtenção

do vocabulário. Assim, nesta etapa foi possível comprovar a efetividade da metodologia e apontar questões que serão apresentadas mais adiante como trabalhos futuros.

Com o intuito de verificar a efetividade do vocabulário para a interação homem-robô, um reconhecedor de gestos teve que ser implementado. Com esse fim, esta etapa consistiu em realizar uma revisão bibliográfica sobre reconhecedores de gestos, objetivando encontrar aquele que melhor se adequasse a cada um dos gestos obtidos. Após todos esses estudos, um reconhecedor foi implementado e testado.

Por fim, após a realização dos estudos, experimentos e de posse de todos os resultados, foi possível a realização deste trabalho escrito.

1.6 Recursos

Para que fosse possível realizar todas as etapas que resultaram neste trabalho, foi necessária a utilização de alguns recursos de *hardware* e de *software*, que podem ser visualizados na listagem a seguir:

1. Para a aquisição dos dados dos voluntários, foi utilizado um sensor do tipo *Microsoft Kinect 360* ([MICROSOFT, 2015](#)), o qual foi cedido pelo GPR-UFS.
2. Para a realização dos experimentos, foi utilizado um robô móvel do tipo Pioneer 3-DX ([ROBOT, 2011](#)), também cedido pelo GPR-UFS.
3. Para a realização dos experimentos e análise dos dados, foi utilizado um computador pessoal com um sistema operacional de 64 *bits*, placa de vídeo com 2 giga *bytes* de memória dedicada, memória principal e secundária com capacidade de armazenamento de 8 giga *bytes* e 1 tera *byte* respectivamente e processador de 2,2 GHz com 4 núcleos físicos e 4 virtuais.
4. Para a implementação do sistema de captura de imagens, foi utilizada a linguagem Java ([GOSLING et al., 2015](#)) e a IDE Netbeans ([ORACLE, 2015c](#)), ambos fornecidos pela *Oracle Corporations* ([ORACLE, 2015a](#)) respectivamente nas suas versões 8.0_45 e 8.0.2.
5. Para analisar os gestos que deram origem ao vocabulário, foi utilizada a linguagem SQL por meio do banco de dados MySQL ([ORACLE, 2015b](#)), também fornecido pela *Oracle Corporations*.

1.7 Contribuições Esperadas

Dentre as contribuições deste trabalho, destacam-se principalmente:

1. A proposta de uma metodologia para a obtenção de vocabulários de gestos intuitivos;
2. A obtenção de um vocabulário de gestos que possa ser utilizado em futuras pesquisas;
3. Contribuição para o desenvolvimento de interfaces intuitivas que facilitem a popularização da robótica.
4. Facilitação de futuras pesquisas na área de reconhecimento de gestos e comparação de resultados, por meio da divulgação no sítio *web* www.gprufs.org dos dados obtidos nos experimentos e nas análises realizadas.

1.8 Estrutura da Dissertação

Esta dissertação está estruturada da seguinte forma:

Capítulo 1 - Introdução: Aqui é dada uma introdução sobre o trabalho, onde é mostrada a motivação que levou à escolha do tema, a apresentação da hipótese de trabalho, solução proposta, as contribuições almejadas e como o mesmo está estruturado.

Capítulo 2 - Estudo dos Gestos e sua Intuitividade: Neste capítulo é apresentada a fundamentação teórica na qual o trabalho se apoiou para tratar dos gestos e sua intuitividade. Para tanto, são apresentados conceitos, classificação e comportamentos dos gestos e os conceitos sobre intuitividade e gestos intuitivos.

Capítulo 3 - HRI Baseada em Gestos: Neste capítulo são abordados os estudos realizados sobre HRI baseadas em gestos, onde é mostrado como os trabalhos abordam os gestos e qual o grau de importância que está sendo dado à questão da intuitividade. Dentro deste mesmo capítulo, são mostradas ainda as metodologias utilizadas para se encontrar gestos intuitivos e quais são os seus pontos fortes e fracos.

Capítulo 4 - Metodologia Proposta: Aqui é descrita a metodologia proposta como uma solução para a problemática estudada. Tal metodologia está dividida em partes bem distintas e explicada de maneira a permitir a sua reprodução.

Capítulo 5 - Aplicação da Metodologia Proposta: Neste capítulo é descrito como a metodologia foi reproduzida e aplicada. Assim, é mostrado como cada uma de

suas etapas foi realizada e quais foram os resultados obtidos. Então, por meio dos resultados obtidos, são discutidas e apontadas as principais vantagens e desvantagens encontradas na metodologia proposta.

Capítulo 6 - Reconhecimento dos Gestos Obtidos no Vocabulário: Aqui é mostrada a implementação de um reconhecedor de gestos, para que assim seja possível verificar se os gestos escolhidos podem ser reconhecidos e consequentemente implantados em robôs.

Capítulo 7 - Conclusões e Trabalho Futuros: Neste último capítulo são apresentadas as conclusões sobre o trabalho e apontados como trabalhos futuros, quais os principais passos a serem dados na sequência deste trabalho.

2 Estudo dos Gestos e sua Intuitividade

2.1 Estudo dos gestos segundo McNeill (1992)

Após mais de 10 anos de estudos e pesquisas, no ano de 1992, David McNeill escreveu seu livro intitulado *"Hand and Mind: What Gestures Reveal About Thought"*. Este trabalho traz uma abordagem muito completa sobre o estudo dos gestos, e por isso foi o livro mais vendido sobre psicolinguística naquele mesmo ano.

Dessa maneira, por meio dele será mostrada a maioria do embasamento teórico a respeito dos gestos no qual este trabalho se sustenta. Assim, as próximas seções abordarão questões como: os tipos de gestos existentes, qual o nível de importância que eles carregam, como eles se distribuem em determinados contextos, como os mesmos são apresentados por diferentes pessoas e outras questões tidas como relevantes para o presente trabalho.

2.1.1 Sobre os Gestos

Quando as pessoas conversam, elas realizam movimentos de maneira espontânea, os quais podem ser chamados de gestos. Tais movimentos são comumente realizados por meio das mãos e braços, e em sua maioria estão sincronizados com o fluxo da fala.

Alguns gestos são executados quase na total ausência da fala, no entanto a maioria deles são realizados no momento em que ela ocorre. Essa dependência é tão forte, que na afasia da fala os gestos são “quebrados” junto com ela, isso porque ambos partem da mesma unidade de criação semântica. No entanto, McNeill deixa claro que, mesmo levando em conta que a fala e os gestos são criados pela mesma fonte semântica, eles não são redundantes, pois, ao serem transmitidos, seguem por canais diferentes, fazendo com que um seja complementar ao outro, o que ele chamou de co-expressividade.

O autor diz ainda que a locução verbal é produzida durante a fase de preparação do gesto. Dessa maneira, um significado presente em um gesto também estará presente na fala. No entanto, enquanto a fala pode esconder os pensamentos, estes têm sua forma global, sintética e imagística mostrada por meio dos gestos espontâneos. Assim, um gesto espontâneo, além de ter sua essência baseada no significado, possui a característica de ser realizado para demonstrar fenômenos reais. Por consequência, os gestos espontâneos são sinteticamente globais, de maneira não combinada, sensíveis ao contexto, de forma

idiossincrática (cada indivíduo segue seu próprio padrão de realização) e possuem tempo de antecipação e de sincronização com a fala.

Por outro lado, existem também os gestos não espontâneos, representados por aqueles que tendem a se assemelhar às linguagens. Isso porque, mesmo ocorrendo ou não durante a fala, estes são executados de maneira pensada, fazendo com que exista, de forma intencional ou não, uma gramática que os integram dentro de uma locução. Logo, enquanto os gestos espontâneos podem ser levados a revelar o estágio primitivo da expressão oral, os sinais convencionais, mesmo quando são gestuais, são o estágio final de suas expressões orais.

2.1.2 Tipos de Gestos

Todos os movimentos visíveis realizados pelo orador são primeiro diferenciados entre gestos e não gestos, sendo que os últimos compreendem auto-toques¹ e manipulação de objetos. O restante é considerado como gestos, e são classificados em um determinado tipo. Além disso, dependendo da capacidade de retratar imagens, um determinado tipo de gesto pode ser denominado, imagístico ou não imagístico

Para McNeill os gestos podem ser classificados em 4 principais categorias: icônicos, metafóricos, dêiticos e batidas.

2.1.2.1 Icônicos

São aqueles que em sua execução, ou na maneira de construção, referem-se a um evento concreto, uma ação, ou mesmo um objeto, que está sendo referenciado pelo discurso ao mesmo tempo em que eles ocorrem. Esse tipo de gesto expressa diretamente uma categoria de significado que deve ser comparado com a fala. Considerando, por exemplo, que no momento do discurso o orador fala algo como “irei **cortar** o papel”, sendo que ao iniciar a palavra cortar, ele realiza um gesto como o ilustrado na Figura 1. É fácil perceber que esse gesto representa a imagem de uma tesoura, um objeto que serve para cortar algo. Dessa maneira, a iconicidade do gesto traz ainda mais significado para a comunicação.

¹ São movimentos que o interlocutor realiza tocando o seu próprio corpo. Por exemplo: Em um processo de interação o interlocutor pode passar uma mão na outra, passar a mão nos cabelos ou braços. Assim, esses movimentos não carregam significado semântico e por isso McNeill não os classifica como gestos.



Figura 1 – Exemplo de um gesto icônico ([PRESS, 2012](#))

2.1.2.2 Metafóricos

Se assemelham aos icônicos, uma vez que apresentam imagens, mas tais imagens possuem um conceito abstrato, tal como conhecimento, linguagem própria, gênero da narrativa, etc..

Os gestos metafóricos são, no entanto, intrinsecamente mais complexos que os icônicos, pois, enquanto um gesto icônico descrever um evento ou objeto concreto por meio da criação de uma homologia aos eventos e objetos, um gesto metafórico deve descrevê-los de uma só vez. Assim, ao representar um objeto, usa-se um gesto icônico (como por exemplo, a mão na forma de copo) e ao representar um evento (como por exemplo, passar por um tubo) também se usa um gesto icônico. Porém, ao fazer os dois de maneira unívoca (empurrar o copo dentro de um tubo) tem-se uma metáfora. Logo, usando esse exemplo, a base do gesto seria o copo e a referência seria “passar pelo tubo”, como ilustra a Figura 2.



Figura 2 – Exemplo de um gesto metafórico ([MCNEILL, 1992](#))

2.1.2.3 Dêiticos (Apontar)

Dêiticos são os movimentos de indicação (apontar para algo), construídos, em sua maioria, por meio do uso do dedo indicador. No entanto, qualquer objeto extensível ou parte do corpo pode ser utilizada, incluindo a cabeça, nariz, ou queixo, bem como objetos manipulados. Os gestos dêiticos que aparecem nas narrativas, raramente ou quase nunca, apontam para algo que não seja uma entidade concreta. Eles selecionam uma parte do espaço, sendo que o significado do gesto depende do valor referencial vinculado a esta região. Um exemplo de um gesto dêitico pode ser visto na Figura 3.

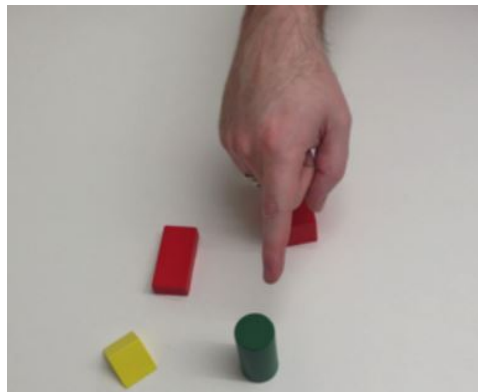


Figura 3 – Exemplo de um gesto dêitico ([MATUSZEK et al., 2014](#))

2.1.2.4 Batidas

Os gestos de batidas podem ser definidos como movimentos que não apresentam significados nítidos, podendo ser reconhecidos positivamente em termos dos seus movimentos típicos e característicos. Eles são, em sua maioria, bifásicos (compostos por apenas dois movimentos), pequenos, de baixa intensidade e possuem um rápido movimento dos dedos e/ou das mãos; carecem de um espaço de gesto especial, e são construídos de fato em qualquer que seja a mão e onde elas se encontrem, incluindo até a posição de descanso. Um exemplo de um gesto de batida, é quando no momento da comunicação o interlocutor pausa a fala e começa a contar algo nos dedos (Figura 4), dessa maneira a contagem segue um certo ritmo e o gesto tem mais significado para o orador do que para o ouvinte.



Figura 4 – Exemplo de um gesto de batida (PAUL; PINKY, 2013)

2.1.2.5 Os gestos emblemáticos

Além das 4 categorias supracitadas, existem ainda gestos que possuem significado próprio sem a necessidade da fala para complementá-lo: estes são os gestos emblemáticos. Apesar de não serem classificados entre as principais categorias, os gestos emblemáticos formam parte de um código social importante, e por isso também devem ser abordados aqui.

Mesmo fazendo parte de um código social, os gestos emblemáticos não são completamente estruturados como uma linguagem. Eles possuem nomes ou normas parafrásicas², que são aprendidas como símbolos específicos e podem ser usados como se fossem palavras faladas. De fato, eles são palavras e/ou frases não faladas, mas não seguem uma gramática e raramente são combinados.

Em contraste com os gestos espontâneos, icônicos e metafóricos, os gestos emblemáticos são fortemente influenciados pela variação regional. Assim, mesmo a forma do gesto sendo a mesma, cada cultura determina o seu significado.

Segundo McNeill, devido à possibilidade de representarem uma alta unidade de significado, os gestos emblemáticos podem ser elementos de uma combinação hierárquica entre eles e os gestos espontâneos.

Os gestos emblemáticos são usados em saudações, para dar comandos, pedir, responder a algum desafio, insultar, proteger, expressar desprezo ou medo. Um exemplo de gesto emblemático pode ser visto na Figura 5, que representa o conhecido gesto de “Ok” utilizado por pessoas de diferentes partes do mundo.

² Normas parafrásicas são aquelas que regem as paráfrases. Uma paráfrase pode ser definida como: Explicação ou tradução mais desenvolvida de um texto por meio de palavras diferentes das nele empregadas. (MICHAELIS, 2009)

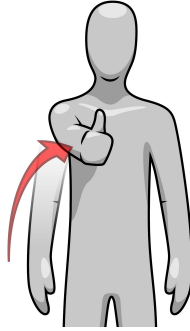


Figura 5 – “Ok”: Gesto emblemático muito utilizado a nível mundial

2.2 Intuitividade dos Gestos

Segundo [Davidson \(1882\)](#) a palavra intuição tem sua origem no latim, *intuitione*, e tem a seguinte etologia: percepção visual, aquilo que se vê por meio dos olhos. Dessa maneira, o autor define a intuição como a percepção imediata de um objeto externo assim que o mesmo é visto. Sendo que a percepção imediata significa a não existência da necessidade do raciocínio prévio para se analisar o objeto em questão. Dessa maneira, um gesto intuitivo pode ser entendido como um gesto que pode ser percebido e interpretado por um determinado indivíduo de maneira imediata, sem que haja a necessidade de uma análise prévia de sua realização.

De acordo com [Wexelblat \(1998\)](#), para ser intuitivo o gesto deve ser realizado da maneira mais natural possível. Isso pode ser comprovado, quando [McNeill \(1992\)](#) explana que os gestos naturais/espontâneos possuem um núcleo de significado maior que os gestos realizados com restrições, como no caso de linguagens de sinais, por exemplo. Como foi visto, [McNeill \(1992\)](#) afirma que os gestos são idiossincráticos, o que significa que uma mesma unidade de significado pode ser representada por diferentes pessoas através de diferentes gestos, uma vez que cada pessoa possui características cinéticas próprias. Dessa maneira, um vocabulário de gestos intuitivos não pode sofrer restrições em sua execução, e deve ser concebido de maneira a maximizar a realização imediata por parte dos indivíduos. Por esse motivo, ao estabelecer que um vocabulário de gestos deve ter apenas um gesto vinculado a cada tarefa, impõe-se uma restrição no vocabulário que pode acabar prejudicando sua intuitividade, pois indivíduos diferentes podem expressar o significado de uma determinada função por meio de gestos diferentes, os quais não deixam de ser intuitivos.

3 HRI baseada em gestos

A coexistência e interação entre homens e robôs é considerada uma das mais importantes questões na área de pesquisa em robótica nos dias atuais (FUJII; LEE; OKAMOTO, 2014). Assim, muitos trabalhos abordam a questão da interação homem-robô, e dentre eles, muitos são dedicados ao caso específico da interação baseada em gestos (BARATTINI; MORAND; ROBERTSON, 2012; WALDHERR et al., 1998; SIEKMANN; HARTMANIS; LEEUWEN, 1999; BROCCIA; LIVESU; SCATENI, 2011; OBAID et al., 2014; MASHOOD; NOURA, 2015).

Segundo Hassani (2011), interfaces baseadas em gestos são mais aceitas pelos usuários que as interfaces táteis. Neste trabalho, o autor busca abordar quais são os tipos de gestos relevantes para serem utilizados em sua pesquisa, usando como referência os tipos de gestos apresentados por Kendon (2004) e McNeill (1992). Nesse contexto, ele informa apenas que os gestos utilizados são do tipo dêitico, mas não menciona se tais gestos foram criados de maneira empírica ou à luz de algum trabalho científico.

Em Waldherr et al. (1998), propõe-se a construção de uma interface de comunicação para HRI baseada em gestos dinâmicos (Figura 6). Eles se propuseram a resolver os problemas de outros trabalhos que abordam a interface gestual apenas com gestos estáticos. Assim, utilizando gestos dinâmicos, os autores afirmam que a interação seria natural. O problema aqui encontrado, é que não foram apresentados quais os critérios/métodos utilizados na escolha dos gestos, o que produz uma dúvida quanto à naturalidade de execução dos mesmos.

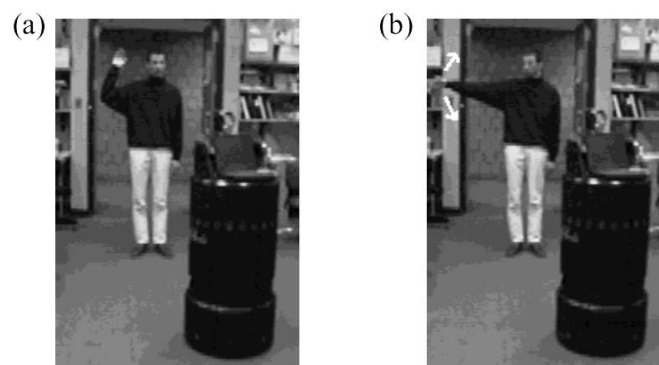


Figura 6 – Gestos apresentados por Waldherr et al. (1998)

Em Meena, Jokinen e Wilcock (2012) é apresentado o desenvolvimento de uma interface de comunicação que emprega fala e gestos manuais para que um robô do tipo NAO possa interagir com as pessoas. Nesse trabalho são utilizados apenas os gestos que

podem ser representados com a palma da mão voltada para cima ou para baixo (Figura 7). Esses gestos foram criados a partir de duas famílias de gestos baseados em “mão aberta” proposta por [Kendon \(2004\)](#). Isso aponta para a possibilidade de que alguns gestos intuitivos podem ser definidos através de estudos psicolinguísticos.

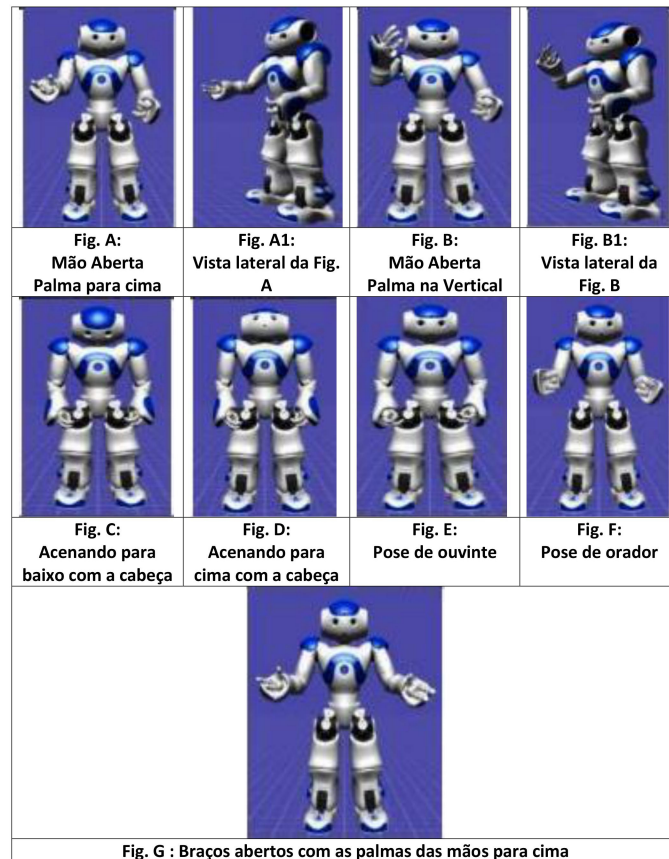


Figura 7 – Gestos apresentados por [Meena, Jokinen e Wilcock \(2012\)](#)

Para [Barattini, Morand e Robertson \(2012\)](#), na indústria automobilística, apesar de haver vários robôs executando as mais diversas tarefas, existem algumas delas pelas quais o ser humano é responsável. Algumas dessas tarefas trazem risco para a saúde e segurança dos trabalhadores, motivo pelo qual a substituição do homem pelo robô, ou ao menos fazer com que ambos trabalhem de forma colaborativa, é desejável. Nesse contexto, uma boa opção consiste no uso de uma interface de comunicação baseada em gestos para que os trabalhadores e robôs possam interagir, já que em ambientes industriais a fala não produz um efeito satisfatório, devido à presença de vários tipos de ruídos emitidos pelas máquinas. Os autores apontam ainda que as linguagens de sinais são ricas em características sintáticas e semânticas e oferecem uma boa interface de comunicação sem a necessidade da fala, mas, por outro lado, elas demandam uma grande carga de trabalho por parte dos usuários

para um devido treinamento, assim tornam-se inviáveis para o uso por trabalhadores em ambientes industriais. Desta maneira, foram escolhidos 19 gestos (Figura 8) a partir da Diretiva Europeia 92/58/EEC (HSE, 2009) que exprimem um determinado comando com menos gestos que uma linguagem de sinais. Além disso, esses gestos foram criados para serem utilizados de maneira que se ofereça mais segurança e entendimento em ambientes industriais, pois enquanto alguém está realizando os gestos, outro, por exemplo, poderá estar no controle de uma máquina. Sendo assim, o trabalho mostra a importância de se ter gestos previamente definidos que possam ser usados em outras pesquisas científicas, ajudando a diminuir o retrabalho por parte daqueles pesquisadores que utilizam gestos como interfaces de comunicação.

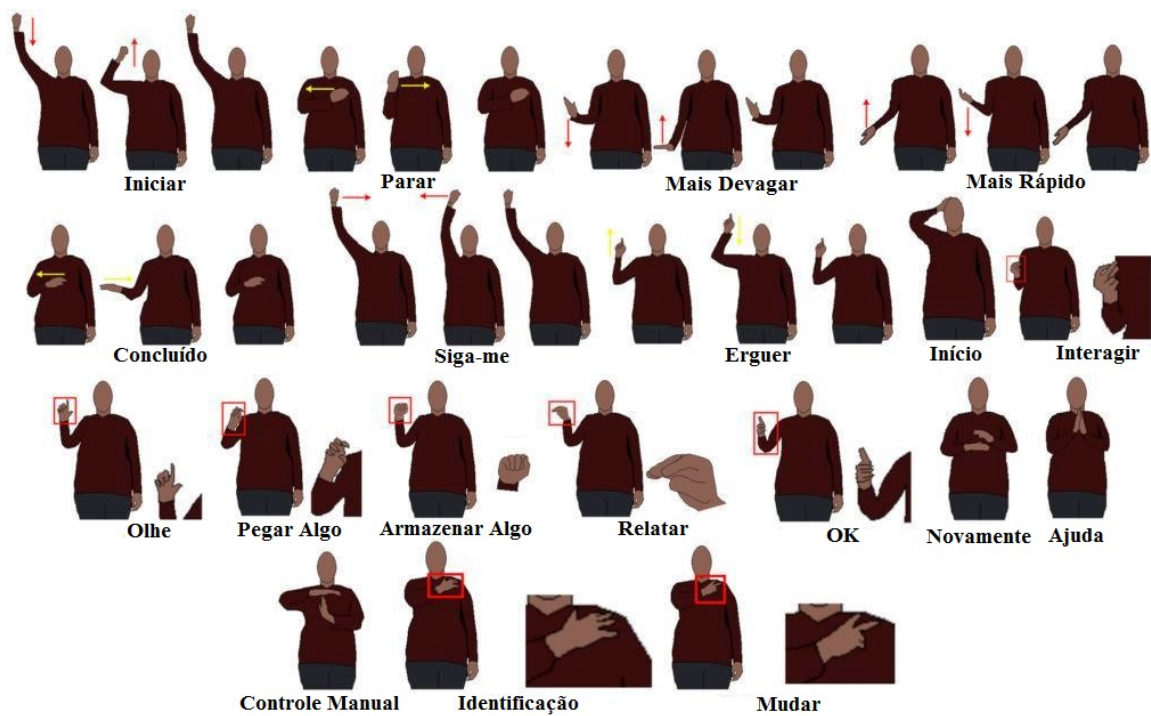


Figura 8 – Gestos apresentados por Barattini, Morand e Robertson (2012)

Como mencionado anteriormente no trabalho de Yohanan e MacLean (2011), foi criado um dicionário de gestos táteis (Figura 9) para tentar resolver o problema que se estabelece ao procurar gestos predefinidos para o uso experimental e/ou industrial. Assim, no decorrer do seu trabalho, o dicionário de gestos foi apresentado, mas não foi mostrado como os gestos que o compõem foram escolhidos. O trabalho traz uma grande contribuição para futuros trabalhos científicos que buscarem estudar comportamentos através de gestos táteis. Porém, assim como em (COSTALONGA et al., 2014), o aspecto da intuitividade não

foi levado em consideração na elaboração desse dicionário, dificultando o emprego de uma interface tátil baseada nesses gestos por pessoas não previamente treinadas.

Gesture label	Gesture definition	Gesture label	Gesture definition
Contact Without Movement	Any undefined form of contact with the Haptic Creature that has no movement. For example: laying one's hand a top the Haptic Creature, or resting one's arm alongside it.	Press	Exert a steady force on the Haptic Creature with your flattened fingers or hand.
Cradle	Hold the Haptic Creature gently and protectively.	Pull	Exert force on the Haptic Creature by taking hold of it in order to move it towards yourself.
Finger Idly	Gently and randomly pull at the hairs of the Haptic Creature's fur with your fingers.	Push	Exert force on the Haptic Creature with your hand in order to move it away from yourself.
Grab	Grasp or seize the Haptic Creature suddenly and roughly.	Rock	Move the Haptic Creature gently to and fro ^a or from side to side.
Hit	Deliver a forcible blow to the Haptic Creature with either a closed fist or the side or back of your hand.	Rub	Move your hand repeatedly to and fro ^a on the fur of the Haptic Creature with firm pressure.
Hold	Grasp, carry, or support the Haptic Creature with your arms or hands.	Scratch	Rub the Haptic Creature with your fingernails.
Hug	Squeeze the Haptic Creature tightly in your arms. Hold the Haptic Creature closely or tightly around or against part of your body.	Shake	Move the Haptic Creature up and down or side to side with rapid, forceful, jerky movements.
Kiss	Touch the Haptic Creature with your lips.	Slap	Quickly and sharply strike the Haptic Creature with your open hand.
Lift	Raise the Haptic Creature to a higher position or level.	Squeeze	Firmly press the Haptic Creature between your fingers or both hands.
Massage	Rub or knead the Haptic Creature with your hands.	Stroke	Move your hand with gentle pressure over the Haptic Creature's fur, often repeatedly.
Nuzzle	Gently rub or push against the Haptic Creature with your nose or mouth.	Swing	Move the Haptic Creature back and forth or from side to side while suspended.
Pat	Gently and quickly touch the Haptic Creature with the flat of your hand.	Tap	Strike the Haptic Creature with a quick light blow or blows using one or more fingers.
Pick	Repeatedly pull at the Haptic Creature with one or more of your fingers.	Tickle	Touch the Haptic Creature with light finger movements.
Pinch	Tightly and sharply grip the Haptic Creature's fur between your fingers and thumb.	Toss	Throw the Haptic Creature lightly, easily, or casually.
Poke	Jab or prod the Haptic Creature with your finger.	Tremble	Shake against the Haptic Creature with a slight rapid motion.

Texto mantido em inglês conforme o original.

Figura 9 – Dicionário de gestos táteis apresentados por Yohanan e MacLean (2011)

Em seu trabalho, Pereira (2012) mostra que é possível que um robô interaja com um ser humano e possa ajudá-lo na realização de tarefas específicas, sendo que para isso, o robô deve possuir um mínimo possível de sensores de variados tipos para que os erros no reconhecimento da interface de comunicação utilizada e no desempenho das tarefas sejam minimizados. Para a comprovação de sua hipótese, Pereira utilizou-se de diversos sensores para detectar os seres humanos e interagir com eles. A interface de comunicação utilizada entre homem e robô foi baseada exclusivamente em gestos, os quais foram reconhecidos através da utilização de um método de segmentação de pele. Para realizar os experimentos, o autor propôs um conjunto de cinco gestos manuais e estáticos (Figura 10), onde cada um representava uma tarefa específica a ser realizada pelo robô. Após realização dos experimentos, pôde-se perceber que um robô pode efetivamente interagir com um ser

humano e ajudá-lo em tarefas específicas. Uma questão a ser considerada, é que mesmo o trabalho tendo se baseado apenas em gestos como interface de interação, como em outros trabalhos citados anteriormente, não houve preocupação em mostrar o porquê da escolha dos cinco gestos utilizados, deixando mais uma vez um questionamento quanto à intuitividade dos mesmos e sua possível utilização em outras pesquisas.



Figura 10 – Gestos apresentados por [Pereira \(2012\)](#)

Para [Burke et al. \(2004\)](#) é de fundamental importância analisar os efeitos da interação homem-robô para os ambientes urbanos de resgate, uma vez que, segundo os autores, é necessário saber o impacto causado pelos robôs na realização do trabalho humano. Sendo assim, após acompanharem algumas situações onde houve a interação entre homem e robô, por meio da gravação em vídeos, eles puderam analisar e perceber que algumas formas de interação são predominantes, sendo a interação não verbal (utilização de gestos) aquela utilizada com maior frequência. Dessa forma, os autores destacaram alguns gestos realizados na interação, os quais podem ser vistos na Tabela 1. Estes gestos, apesar de não terem sido propostos como um vocabulário, trazem uma ideia de algumas funções que se usam na interação entre homens e robôs móveis em um ambiente de resgate.

Tabela 1 – Tabela do gestos estudados em [Burke et al. \(2004\)](#) (tradução própria)

Gestos	Descrição
Vir para frente	Movimento na direção do robô para que ele se movimente para frente
Polegar para cima	Mão fechada no punho e indicador esticado para cima
Parar	Mão parada e esticada na direção do robô
Apontar	Usando os dedos para apontar para um lugar ou objeto específico
Sinal de OK	Aproximando o polegar e o indicador formando um círculo
Outros	Outros gestos (normalmente conversacionais, no entanto sem intenção de destinar uma mensagem)

Em [Droeschel et al. \(2012\)](#) é apresentado um robô doméstico chamado de “Cosero” o qual foi utilizado para competir na *Robot Cup Home* em 2011 ([ROBOTCUP, 2011](#)). O robô apresentado ajuda seres humanos a realizarem tarefas cotidianas. Assim, os autores

descrevem que ele foi dotado de capacidade de interação não verbal e por isso era capaz de interagir com seres humanos através de um conjunto de gestos predefinidos. Apesar do trabalho ser muito interessante quanto a sua objetividade, ele poderia ser melhor aproveitado caso os autores tivessem a preocupação de descrever os gestos que o robô utilizava para interagir e como estes foram escolhidos, pois assim, outros trabalhos poderiam usufruir dos gestos citados ou ainda utilizar a mesma metodologia para buscar gestos que correspondessem a outras funções. Além do mais, pelo fato de o robô se dedicar a ajudar pessoas em tarefas domésticas cotidianas, os gestos devem ser os mais intuitivos possíveis, o que não pode ser inferido apenas com as informações fornecidas no trabalho.

No trabalho apresentado por [Fujii, Lee e Okamoto \(2014\)](#), é mostrado um estudo sobre o reconhecimento de gestos dinâmicos baseados em visão RGB-D (imagem colorida acrescida de suas informações de profundidade - do inglês *Red, Green, Blue and Depth*). Os autores testaram o reconhecedor em um ambiente real, onde o robô teve que interagir com pessoas para receber comandos que o fizessem executar algumas tarefas domésticas. Para isso, a interação foi baseada totalmente em gestos, os quais eram realizados pelos usuários e reconhecidos pelo robô por meio do reconhecedor proposto. Mesmo em se tratando de um trabalho que foca apenas no reconhecimento dos gestos, e não na obtenção de vocabulários, ele mostra quais foram os gestos utilizados em seus experimentos (Figura 11).

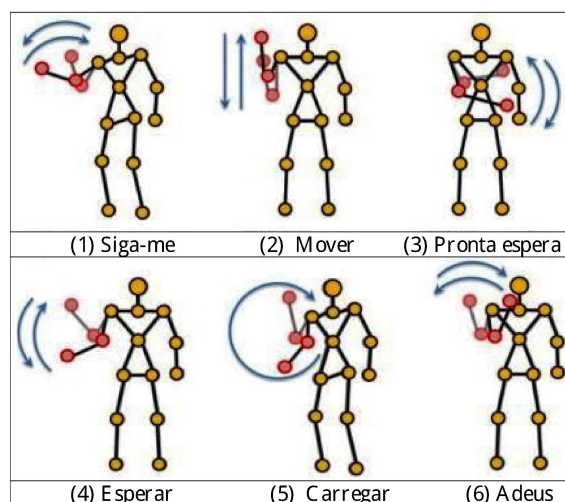


Figura 11 – Gestos apresentados por [Fujii, Lee e Okamoto \(2014\)](#)

No trabalho de [Yang, Park e Lee \(2007\)](#) é abordada a questão da importância do reconhecimento de gestos de corpo inteiro. Nele os autores apontam para a questão de que muitos gestos não são puramente manuais e que para se ter uma interação mais efetiva,

deve-se levar em consideração os gestos realizados por todos os membros do corpo. Os autores apresentam os gestos utilizados em seus estudos (Figura 12), mas não mostram como os mesmos foram obtidos ou de onde foram extraídos. O mesmo se verifica nos trabalhos de [Ohta \(2014\)](#), [Lee, Cha e Kwon \(2014\)](#), [Kim, Nam e Park \(2014\)](#).

Gestos	Exemplos de Gestos	Gestos	Exemplos de Gestos
Caminhando em um lugar		Correndo em um lugar	
Dobrando na cintura		Pulando em um lugar	
Deitando no chão		Acenando com a mão	
Sentando no chão		Elevando a mão direita	
Descendo até o chão		Tocando o joelho e a cintura	

Figura 12 – Gestos apresentados por [Yang, Park e Lee \(2007\)](#)

Dessa forma, mesmo que muitos trabalhos procurem abordar os gestos como interface de comunicação para HRI, não se encontram muitos trabalhos que apresentem o conjunto de gestos utilizados, sendo que aqueles que os apresentam, não mostram como tais gestos foram escolhidos, muito menos se foi utilizado algum embasamento científico na escolha dos mesmos, além de não abordar a questão a respeito da intuitividade dos gestos utilizados.

Para [Wexelblat \(1998\)](#), se a intuitividade não for levada em consideração, não faz sentido utilizar interfaces naturais, uma vez que os usuários terão que aprender os comandos a serem dados como parâmetros de entrada para a interface, quando, já que a intuitividade não é o foco da interface, poderiam ser utilizadas interfaces mais simples e efetivas, como as baseadas em comandos manuais (*joytick* e/ou controle remoto por exemplo), que possuem uma taxa de acerto muito alta e uma baixa complexidade de

implementação. Logo, um primeiro passo a ser dado neste sentido, seria a utilização de uma metodologia que permitisse a obtenção de gestos intuitivos.

3.1 Metodologias para obtenção de vocabulários de gestos

3.1.1 Metodologia proposta por Wachs

Percebendo a importância da intuitividade do vocabulário de gestos e também o problema de se ter uma metodologia eficiente para a busca de vocabulários de gestos para HRI, Wachs (2006) resolveu buscar tal metodologia na área de HCI. Essa foi uma decisão acertada, uma vez que segundo Powers (2008) a área de HCI, por ser uma área com maior desenvolvimento, pode oferecer soluções para problemas ainda não resolvidos para HRI. Assim, Wachs apresenta uma abordagem sobre os vocabulários de gestos para HRI baseados nos estudos psicolinguísticos (McNeill, 1992) e em estudos sobre HCI (Nielsen et al., 2004). Neste trabalho, ele trás uma definição para intuitividade, uma metodologia sintética para a escolha de um vocabulário “ótimo” e a descrição da aplicação de tal metodologia. A metodologia apresentada faz uso dos conhecimentos fornecidos pela metodologia apresentada por Nielsen et al. (2004), a qual será referida no presente trabalho apenas como “metodologia de Nielsen”, para buscar maximizar a intuitividade dos gestos do vocabulário. Apesar de ser o único trabalho encontrado em HRI que descreve os gestos estudados (Figura 13) juntamente com uma metodologia clara para a escolha dos mesmos, a metodologia empregada leva em consideração apenas gestos estáticos e um vocabulário onde a relação entre gestos e atividades é de um para um. Dessa maneira, quando se restringe a interface a responder apenas a gestos estáticos, deixa-se de lado a naturalidade de realização dos mesmos. Outro ponto importante, é que quando a atividade é condicionada a ser representada apenas por um gesto, faz-se com que os usuários sejam obrigados a realizar apenas aquele gesto para aquela atividade. No entanto, como, segundo McNeill (1992), os gestos são idiossincráticos, cada usuário pode realizar gestos diferentes para a mesma atividade. Sendo assim, ao ser posto em um ambiente real de interação, a taxa de erro do reconhecedor provavelmente aumentaria, fazendo com que a robustez do vocabulário fosse diminuída.

Esse mesmo questionamento também foi abordado pelo autor, que propõe resolvê-lo futuramente por meio de um vocabulário onde exista mais de um gesto para cada

atividade e onde um gesto possa representar mais de uma atividade, o que ele denomina de “vocabulário complexo”. Outro aspecto também importante é que, mesmo o trabalho citando a psicolinguística, o estudo da intuitividade foi feito baseando-se na engenharia de usabilidade abordada por [Nielsen et al. \(2004\)](#), o que mostra que essa questão não foi trabalhada da forma mais adequada.

Dessa maneira, mesmo em se tratando de uma proposta para encontrar um vocabulário de gestos ótimo do ponto de vista psicofisiológico, não existe embasamento científico coerente para o estudo da intuitividade. Esta falta de embasamento pode ser comprovada quando o autor faz uma crítica à metodologia de Nielsen, ao dizer que em seu trabalho, [Nielsen et al. \(2004\)](#) esquece dos aspectos técnicos do vocabulário, pois se baseia apenas nos aspectos humanos (abordagem centrada no usuário). Essa crítica pode ser aceita caso exista a necessidade de incluir no vocabulário os aspectos técnicos referentes à máquina. No entanto, vale salientar que o objetivo de uma interface intuitiva é que os seres humanos possam selecionar comandos de maneira a não necessitem de um intenso treinamento prévio para poderem interagir. Porém, gestos considerados de fácil reconhecimento para a máquina, podem não ser intuitivos para o usuário, o que refuta a possibilidade de inclusão de outros aspectos que não seja o da intuitividade na escolha de vocabulários intuitivos.



Figura 13 – Gestos apresentados por [Wachs \(2006\)](#)

3.1.2 A metodologia de Nielsen

Segundo [Nielsen et al. \(2004\)](#), existem duas abordagens para se obter vocabulários de gestos: a abordagem centrada no usuário e a abordagem centrada na máquina. Na abordagem centrada no usuário, deve-se levar em consideração apenas a intuitividade dos gestos, já na abordagem centrada na máquina, deve-se levar em consideração as características dos gestos que minimizem a taxa de erro do reconhecedor. Desta maneira, objetivando mais efetividade na interação, a metodologia de Nielsen baseia-se na abordagem

centrada no usuário para construir vocabulários de gestos intuitivos. Nela, os autores tentaram buscar um equilíbrio entre a intuitividade e a ergonomia do vocabulário, o que segundo eles pode ser conseguido seguindo os passos descritos a seguir:

Passo 1. Escolha das Funções: Aqui devem ser definidas quais são as funções que o vocabulário de gestos deve representar. Estas funções correspondem aos comandos que a máquina deve receber como resultado da interação por meio dos gestos.

Passo 2. Coleta dos gestos no domínio do usuário: Através de experimentos com voluntários, deve-se coletar dados gestuais (vídeos por exemplo) que representem cada uma das funções a serem realizadas pelo robô. No trabalho, não foi especificado como os experimentos devem ser realizados, no entanto os autores se basearam na abordagem “Mágico de Oz” (do inglês *Wizard of Oz - WoZ*)¹ (DAHLBÄCK; JÖNSSON; AHRENBORG, 1993) para a coleta dos gestos. Um destaque a ser feito, é que cada voluntário realizou apenas um gesto para cada uma das funções, o que permitiu que cada voluntário pudesse realizar tantos experimentos quanto a quantidade de funções escolhidas.

Passo 3. Extração do vocabulário de gestos: Utilizando-se a frequência dos gestos coletados na etapa anterior, escolhem-se os gestos que corresponderão a cada atividade. Aqui os autores escolhem o gesto realizado com maior frequência em cada função.

Passo 4. Avaliação do vocabulário: Para medir a intuitividade e ergonomia de cada um dos gestos escolhidos, as próximas etapas são necessárias:

Semântica: Define a intuitividade dos gestos com base nos erros de associação. Por meio de voluntários que não participaram da etapa de coleta dos gestos, faz-se com que cada um deles associe os gestos escolhidos na etapa anterior a cada uma das funções. Após o resultado, calcula-se o percentual de voluntários que associaram o gesto escolhido com a tarefa à qual ele corresponde.

Memória: Mede o grau de memorização do vocabulário através do erro de execução dos gestos. Para isso, primeiro os gestos obtidos na etapa de captura são vistos por voluntários que não participam dos experimentos anteriores, depois cada

¹ Essa abordagem consiste em fazer o usuário pensar que a máquina está respondendo aos seus comandos, no entanto o que realmente está acontecendo é que a máquina está sendo controlada remotamente por uma outra pessoa. Dessa maneira, como o usuário pensa estar realmente interagindo com a máquina, ele acaba realizando os gestos de maneira mais natural.

um deve executar o gesto vendo apenas a função que ele representa. No cálculo são contabilizados os acertos, os erros e também o tempo que o voluntário levou pensando no gesto. Logo, o objetivo deste procedimento é medir o quanto memorizável é o vocabulário, o que juntamente com a medida da associação semântica, compõem o nível de intuitividade do mesmo.

Esforço: Mede o quanto estressante é o vocabulário através da opinião do voluntário após executar um determinado gesto (200/tamanho_vocabulário) vezes. Dessa maneira, após a realização dos gestos, o voluntário atribui uma característica predefinida para o gesto executado. Cada resposta serve como parâmetro para a medição da ergonomia do vocabulário, sendo elas as seguintes:

1. Sem problemas
2. Ligeiramente Cansativo
3. Cansativo
4. Muito Irritante
5. Impossível

O trabalho não diz como alguns passos anteriores devem ser executados. Dessa maneira, a depender da execução, a metodologia possivelmente trará vocabulários diferentes para o mesmo conjunto de tarefas e o mesmo grupo de voluntários. Outra característica a ser observada é o fato da metodologia ter o seu suporte teórico baseado apenas na engenharia de usabilidade, o que não traz muita segurança quanto ao aspecto intuitivo dos gestos. Deve-se observar ainda que ao utilizar a ergonomia como um dos fatores de decisão para escolha do vocabulário, a intuitividade não se torna a característica principal do vocabulário. Apesar disso, a metodologia de Nielsen foi a mais clara e reprodutível abordagem encontrada para se obter gestos intuitivos, e já foi usada por outros pesquisadores como em [Wobbrock, Morris e Wilson \(2009\)](#) e [Wachs \(2006\)](#).

3.1.3 Outras Metodologias

[Wu, Wang e Zhang \(2016\)](#) apresentam uma metodologia para obtenção de vocabulários de gestos intuitivos. Para isso, os autores baseiam-se em uma abordagem centrada no usuário para fazer a captura dos dados gestuais. Todas as fases nela descritas são iguais às de [Nielsen et al. \(2004\)](#), às quais os autores fazem as devidas referências. No

entanto, a metodologia apresentada trás um diferencial que, segundo os autores, pode gerar vocabulários de gestos com maior representatividade do que aqueles obtidos por meio da aplicação da metodologia apresentada em (NIELSEN et al., 2004). Tal diferencial corresponde a uma mudança na fase de análise dos dados gestuais, onde, enquanto em Nielsen et al. escolhem-se apenas os gestos com maior frequência, de maneira que se tenha apenas um conjunto de gestos representando as funções, a metodologia em questão escolhe os n gestos de maior frequência, obtendo assim vários conjuntos de gestos para representar as funções. Dessa forma, por meio da aplicação da etapa de Avaliação do Vocabulário (que consiste na mesma etapa descrita no passo 4 da metodologia de Nielsen et al. apresentada na seção 3.1.2), dentre os gestos que compõem cada um dos conjuntos, escolhem-se aqueles melhor avaliados para compor o vocabulário resultante.

Apesar de ter realmente um aspecto relevante que a diferencia da metodologia de Nielsen et al., ao fim da sua aplicação, a metodologia descrita acaba gerando um vocabulário simples (um gesto representa apenas uma função e uma função é representada por apenas um gesto) tal como aqueles obtidos com a aplicação das metodologias de Wachs (2006) e Nielsen et al. (2004). Com isso, além de não trazer estudos adequados sobre a intuitividade dos gestos, a metodologia ainda não resolve a questão dos vocabulários complexos deixados por Wachs (2006). Para finalizar, outro problema encontrado na mesma, é que ela não deixa claro como determinar o valor de n , o que pode gerar dúvidas na hora de aplicá-la.

Em seu trabalho, Medeiros, Tavares e Fonseca (2015) tentam responder a seguinte pergunta: como projetar uma interface de usuário baseada em gestos? Para isso, os autores explanam que um passo a ser dado nesta direção é o de encontrar vocabulários de gestos intuitivos. Para isso, o trabalho descreve uma metodologia de obtenção de vocabulários de gestos intuitivos que, por ser fortemente baseada na metodologia apresentada por Nielsen et al. (2004), traz as suas mesmas etapas. No entanto, a fim de melhorar a representatividade dos gestos obtidos, a metodologia descrita acrescenta uma possível iteração entre as etapas de captura e análise dos gestos (Figura 14), sendo esta iteração, a única característica encontrada que diferencia a metodologia descrita da metodologia apresentada por Nielsen et al. (2004).

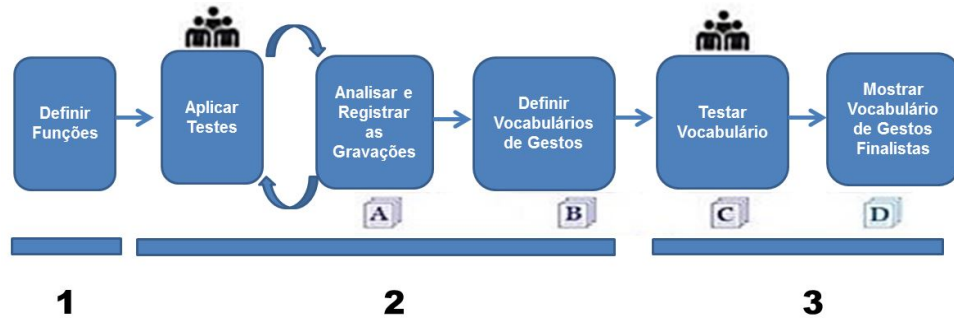


Figura 14 – Etapas de realização da metodologia descrita em [Medeiros, Tavares e Fonseca \(2015\)](#)

3.2 Considerações

De fato, a metodologia de Nielsen foi a abordagem mais plausível encontrada para se obter um vocabulário gestual que leva em consideração o aspecto da intuitividade. No entanto, em [Nielsen et al. \(2004\)](#) as etapas necessárias para se aplicar a metodologia não são completamente explanadas, o que acaba dificultando o seu uso. Além disso, por levar em consideração apenas os conhecimentos da engenharia de usabilidade na concepção da metodologia, a intuitividade acaba não sendo o principal aspecto a ser considerado na escolha do vocabulário, o que não deveria ocorrer, uma vez que o que se busca para um vocabulário gestual é a capacidade de as pessoas realizarem os gestos de maneira imediata, sem a necessidade de uma apresentação prévia. Finalmente, o fato de [Nielsen et al. \(2004\)](#) ignorar completamente a abordagem dos estudos psicolinguísticos consiste em uma falha grave, o que se agrava pelo fato da intuitividade ser uma característica importante a ser considerada.

Portanto, é necessário que se tenha uma metodologia que permita encontrar vocabulários de gestos mais robustos e que sejam intuitivos para os possíveis usuários. Com isso, a comunicação entre homem e robô poderá ser feita de maneira mais natural, o que possibilitará um aproveitamento mais eficaz do potencial que a robótica pode proporcionar para a vida das pessoas.

4 Metodologia Proposta

A metodologia proposta neste trabalho para determinação dos gestos que devem representar um vocabulário intuitivo deve ser realizada por meio de quatro etapas principais: escolha das funções a serem realizadas pela máquina, captura de dados gestuais que representem cada uma das funções, análise dos dados capturados e escolha dos gestos que melhor representem cada uma das funções.

4.1 Escolha das funções

Esta etapa é bem semelhante àquela abordada na metodologia sugerida por [Nielsen et al. \(2004\)](#), onde se deve mapear um conjunto de funções às quais o vocabulário de gestos deve fazer referência. Esta etapa é de fundamental importância, uma vez que as próximas etapas dependerão diretamente de quantas e quais funções foram escolhidas.

4.2 Captura dos dados gestuais

Segundo [Nielsen et al. \(2004\)](#), uma vez que se resolve colher os gestos no domínio do usuário, esse procedimento pode ser realizado por meio de duas abordagens, consciente e subconsciente. Na abordagem consciente o voluntário é convidado a indicar ou ilustrar qual gesto deve representar uma determinada função, ou então pode escolher dentre vários gestos qual é aquele que melhor representa a função em questão. Já na abordagem subconsciente, o voluntário é convidado a interagir com a máquina de maneira a fazê-la executar uma determinada função através de uma interface baseada apenas em gestos. Desta maneira, o voluntário é levado a expressar a sua real percepção de qual gesto corresponde àquela função, ou seja, não existe a necessidade de um intervalo de tempo para pensar sobre o que deve ser feito, uma vez que os gestos fluem diretamente do seu subconsciente. Com isso, podem ser obtidos gestos mais espontâneos, o que segundo [McNeill \(1992\)](#), correspondem a gestos com maior grau de significado.

Como o que se deseja com este trabalho é fornecer um procedimento para se encontrar um conjunto de gestos intuitivos, deve-se utilizar a abordagem subconsciente, executando-a por meio de experimentos que levem o voluntário a realizar gestos da maneira mais espontânea possível. Para isso, a abordagem a ser utilizada é a do “Mágico de Oz”

(do inglês *Wizard of Oz* - *WoZ*) ([DAHLBÄCK; JÖNSSON; AHRENBORG, 1993](#)), onde o voluntário interage com a máquina de modo a pensar que ela está interagindo com ele de maneira autônoma, no entanto, a máquina, sem que o voluntário perceba, está sendo controlada remotamente por uma outra pessoa. Desta maneira, como o voluntário pensa estar realmente interagindo com a máquina, a interação fica mais espontânea e o voluntário tende a realizar os gestos que, subconscientemente, são os mais intuitivos para ele. Os trabalhos de [Nielsen et al. \(2004\)](#) e [Wachs \(2006\)](#), descritos no Capítulo 3, são exemplos de utilização da abordagem *WoZ* para captura de gestos.

Por outro lado, de acordo com [Choi et al. \(2014\)](#), cada voluntário possui em sua mente um conjunto de gestos que representam cada uma das funções a ser realizadas pela máquina. Assim, a depender de quais gestos desse conjunto são capturados nos experimentos, o vocabulário obtido pode ser mais ou menos intuitivo. Dessa maneira, a mera utilização da abordagem *WoZ* não é suficiente para garantir a captura de um conjunto adequado de gestos intuitivos. Para verificar essa questão, [Choi et al. \(2014\)](#) realizaram um conjunto de experimentos tendo como base as metodologias descritas em [Nielsen et al. \(2004\)](#) e [Wachs \(2006\)](#), nas quais os voluntários são instruídos a realizar apenas um gesto para cada função.

Assim, em ([CHOI et al., 2014](#)) os autores realizaram inicialmente um experimento seguindo as especificações descritas em ([NIELSEN et al., 2004](#)) e ([WACHS, 2006](#)), onde foi possível obter um vocabulário de gestos para um determinado conjunto de funções. Um mês depois, foi realizado um outro experimento, utilizando os mesmos participantes. O objetivo principal deste segundo experimento foi o de verificar se cada voluntário, ao escolher novamente um gesto para cada uma das funções, mantinha o gesto que havia realizado no primeiro experimento. Para isso, cada voluntário, antes de escolher um gesto para representar uma função, observava todos os gestos correspondentes àquela função que foram realizados pelos outros voluntários no primeiro experimento.

Após análise dos resultados obtidos, percebeu-se que 66% dos voluntários escolheram um gesto diferente no segundo experimento, e que 70% dos gestos do segundo vocabulário diferiam daqueles obtidos no primeiro experimento. Dessa forma, para validar os resultados obtidos, os autores repetiram o segundo experimento, utilizando desta vez participantes diferentes daqueles que participaram dos experimentos anteriores. Após análise dos novos resultados, os autores verificaram que a maioria dos gestos do terceiro vocabulário era igual aos gestos do segundo. Isso levou os autores a concluir que as metodologias que

usam como único critério de seleção a frequência de repetição de gestos realizados em apenas um conjunto de experimentos, não garantem, necessariamente, a “popularidade” dos mesmos (CHOI et al., 2014).

Dessa maneira, qual seria a abordagem a ser seguida para poder coletar, no domínio do usuário por meio da abordagem WoZ, o conjunto de gestos mais intuitivos para cada voluntário que pode representar uma determinada tarefa? A resposta para esse questionamento não foi dada por Choi et al. (2014).

Como em um único experimento não se tem garantia de que os gestos capturados são os mais representativos, ao invés de convidar o voluntário a realizar o mesmo experimento várias vezes para cada função, seria mais simples capturar em um único experimento, todos os gestos que o voluntário creia que possam ser adequados para a uma mesma função. Assim, propõe-se como solução à questão levantada, que cada voluntário realize gestos correspondentes à função desejada até que suas ideias se esgotem. Para isso, se faz necessário que a interface de comunicação nunca permita que o robô obedeça ao comando do voluntário, e que este seja instruído a realizar gestos até que a interface responda à interação adequadamente. Com isso, cada voluntário realizará todos os gestos que pensa serem intuitivos para que o robô realize uma determinada função, até chegar ao ponto de se frustrar por não haver conseguido interagir, o que será denominado de ***Abordagem Baseada em Frustração (ABF)***. Dessa forma, mesmo o experimento sendo executado apenas uma vez para cada voluntário, ainda assim se obterá um repertório de gestos correspondentes a cada função que possuem um nível de intuitividade relativamente alto, uma vez que são gestos espontâneos.

É importante salientar que cada voluntário deve realizar apenas um experimento para uma única função, pois, uma vez que a frustração for alcançada, o mesmo poderá realizar gestos mais elaborados ou mesmo muito diferentes daqueles que ele pensa serem intuitivos para outra função, já que ele sabe que o robô não responde da maneira satisfatória ao que ele considera ser ideal.

Todos os experimentos devem ser filmados e salvos, para que nas próximas etapas seja possível analisar os gestos realizados e escolher aqueles que devem compor o vocabulário. É importante ainda que cada vídeo seja nomeado por função e numerado com um identificador do voluntário (ID) ao qual ele corresponde, para que assim, posteriormente, seja possível identificar a que função o vídeo está associado e vinculá-lo diretamente ao seu respectivo voluntário.

4.3 Análise dos gestos capturados

Após a coleta dos dados em formato de vídeo, para saber quais gestos deverão fazer parte do vocabulário, necessita-se seguir algumas etapas para analisar os gestos gravados, que serão descritas na sequência.

4.3.1 Análise dos gestos presentes em cada um dos vídeos

Nesta etapa os vídeos precisam ser analisados a fim de se identificar cada um dos gestos realizados. Para isso, se faz necessária a criação de duas tabelas, uma denominada Tabela de Descrição dos Gestos (TDG) contendo uma coluna para guardar o identificador do gesto e outra coluna para a sua descrição, e outra denominada Tabela de Gestos Realizados (TGR) contendo quatro colunas, para receber respectivamente a função correspondente ao gesto, o identificador do voluntário que o realizou, o identificador do gesto (obtido a partir da tabela 1) e o tempo no qual sua realização se iniciou. Assim, a análise dos gestos pode ser realizada seguindo os seguintes passos:

Passo 1: A cada novo gesto realizado, caso o avaliador (pessoa responsável por analisar os gestos capturados) julgue que a realização deste é compatível com algum gesto descrito antes na tabela 1, passa-se para o Passo 3, caso contrário passa-se para o passo 2.

Passo 2: Já que o gesto realizado não corresponde a nenhum gesto descrito, uma nova linha deve ser criada na tabela 1, a qual receberá respectivamente o número que identifica o gesto e a descrição da realização do mesmo.

Passo 3: Após a identificação do gesto, a tabela 2 deve receber o nome da função à qual o gesto pertence, o identificador do voluntário que realizou o gesto, o identificador do gesto (obtido na tabela 1) e o tempo (em segundos) no qual sua realização se iniciou.

Passo 4: Retorna ao Passo 1 até que se tenha analisado todos os gestos de todos os voluntários. Esta subetapa, por ser quase que totalmente subjetiva, deve ser realizada com muita cautela, uma vez que os resultados aqui obtidos influenciarão diretamente na escolha do vocabulário.

4.3.2 Cálculos das taxas de ocorrência

Com o intuito de objetivar ao máximo o processo de escolha dos gestos que deverão compor o vocabulário, propõe-se criar quatro diferentes taxas de ocorrência para cada um dos gestos pertencentes a cada uma das funções.

1. **Taxa de Ocorrência Geral (TOG):** Calcula a proporção de cada gesto dentre todos aqueles realizados na função (levam-se em consideração todos os gestos realizados por todos os voluntários da função). Com isso, será possível medir a importância global de cada gesto em relação a todos os outros gestos realizados em uma determinada função.
2. **Taxa de Primeira Ocorrência (TPO):** Na subetapa anterior, para o cálculo da taxa de ocorrência de cada gesto, foram levados em consideração todos os gestos realizados por cada voluntário. Porém muitas vezes um mesmo voluntário pode tender a realizar o mesmo gesto várias vezes durante a realização do experimento, o que pode levar o gesto a ter um valor de ocorrência alto, mesmo sendo realizado por uma pequena quantidade de voluntários. Desta maneira, deve-se também calcular o valor de ocorrência de cada gesto levando em consideração apenas a sua primeira realização para cada voluntário. Para isso, cada gesto deve ser contabilizado apenas uma vez por voluntário, dividindo-se depois o quantitativo dos voluntários que realizaram tal gesto pela soma de todos os diferentes gestos realizados por todos os voluntários.
3. **Taxa de Ocorrência por Voluntário (TOV):** Mesmo após o cálculo das duas taxas anteriores, ainda não se sabe o percentual de voluntários que realizaram um determinado gesto. Por isso, os gestos devem ser contabilizados apenas uma vez por cada voluntário (como realizado no TPO), no entanto a proporção deve ser calculada em relação à quantidade de voluntários que realizaram os experimentos e não em relação à quantidade dos diferentes gestos por eles realizados, o que permite saber qual percentual de voluntários que realizou cada gesto.
4. **Taxa de Ocorrência por Tempo (TOT):** Pensando na questão de que os gestos realizados primeiro devem ser mais intuitivos, se faz necessário atribuir pesos diferentes para cada gesto de maneira que os gestos realizados primeiro tenham um peso maior. Para poder obter um peso mais justo em relação à ocorrência de cada gesto, pode-se tomar como base o tempo no qual o gesto foi realizado. Assim, para

garantir que os primeiros gestos tenham pesos maiores, o peso de cada gesto deve receber a diferença entre o tempo do último gesto realizado no atual experimento e o tempo de realização do gesto em questão.

Para uma descrição mais clara, considere um modelo genérico, como o representado na Figura 15, onde:

- F é um conjunto de W funções.
- F_i é a i -ésima função do conjunto F , $\forall i = 1, \dots, W$.
- E_i é o conjunto de X_i experimentos realizados para a função F_i . Aqui é importante destacar que uma vez que cada voluntário realiza apenas um experimento, a quantidade voluntários também é X_i .
- E_{i_j} corresponde ao j -ésimo experimento realizado para a função F_i .
- G é o conjunto com N gestos diferentes realizados em F .
- GE_i é o conjunto com M_i gestos realizados nos X_i experimentos da função F_i , tal que $GE_i \subseteq G$.
- GDE_i é o conjunto com N_i diferentes gestos realizados nos experimentos do conjunto E_i , tal que $GDE_i \subseteq G$.
- GE_{i_j} é um conjunto com Z_{i_j} gestos realizados no experimento E_{i_j} , tal que $GE_{i_j} \subseteq GE_i$.
- GDE_{i_j} é um conjunto com Y_{i_j} diferentes gestos realizados no experimento E_{i_j} , tal que $GDE_{i_j} \subseteq GE_{i_j}$.
- $GDE_{i_{j_k}}$ corresponde ao k -ésimo gesto diferente realizado no j -ésimo experimento para na função F_i , tal que $GDE_{i_{j_k}} \subseteq GDE_{i_j}$.
- $GE_{i_{j_k}}$ corresponde ao k -ésimo gesto realizado no j -ésimo experimento para na função F_i , tal que $GE_{i_{j_k}} \subseteq GDF_i$.
- $TE_{i_{j_k}}$ corresponde ao tempo de realização do k -ésimo gesto no j -ésimo experimento realizado para a função F_i .

Os valores das taxas de ocorrência para um gesto qualquer G_{n_i} pertencente a uma função F_i , onde $n_i = 1, \dots, Y_i$ podem ser calculados da seguinte maneira:

1. TOG de um gesto G_{n_i}

$$TOG_{G_{n_i}} = \frac{\text{Somatório de cada gesto } G_{n_i} \text{ realizado na função } F_i}{\text{Somatório de todos os gestos realizados na função } F_i}$$

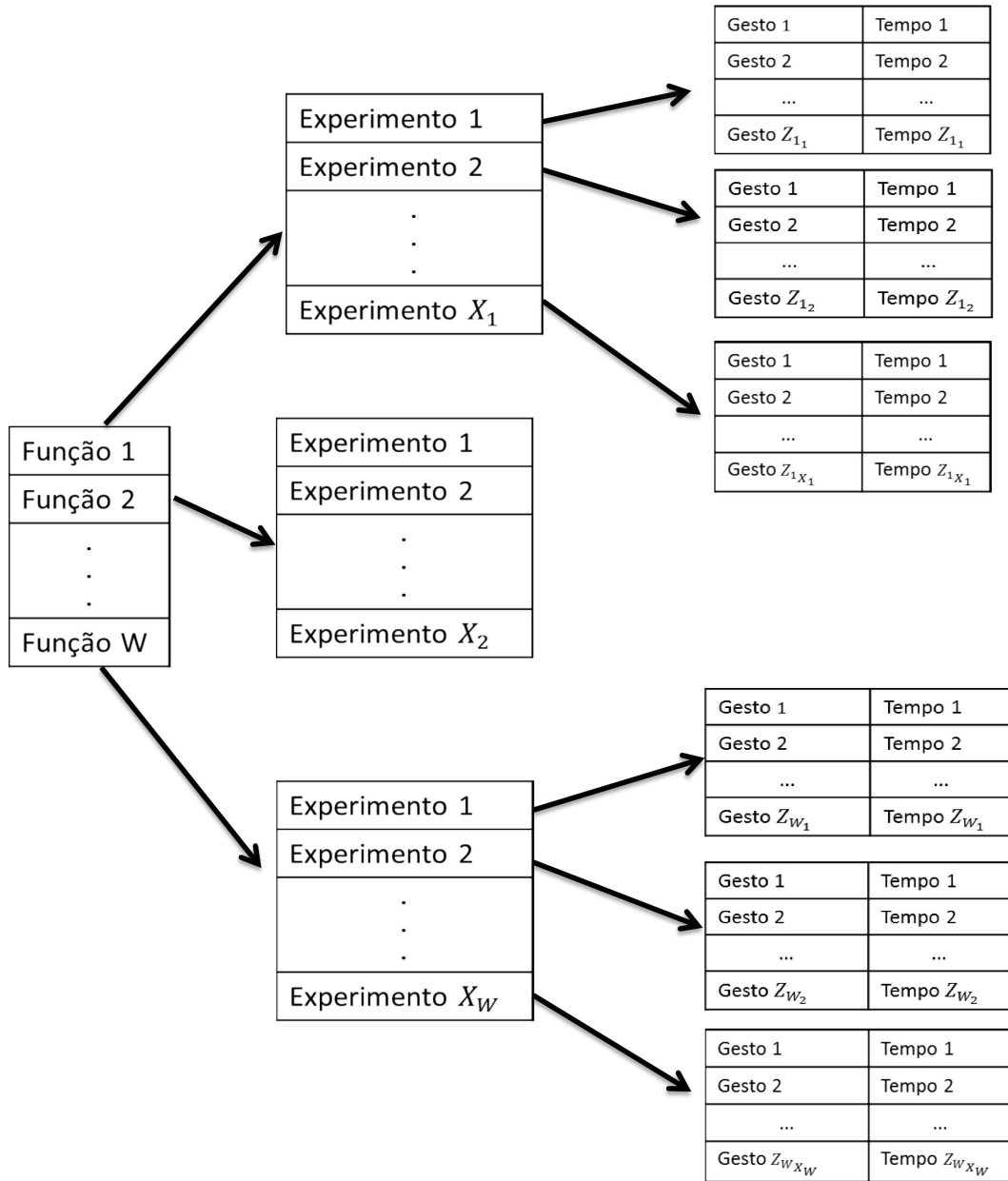


Figura 15 – Modelo genérico do ambiente obtido após a análise dos gestos presentes em cada um dos vídeos

$$TOG_{G_{n_i}} = \frac{\sum_{x=1}^{X_i} \sum_{z=1}^{Z_{ix}} f_1(x, z)}{M_i} \quad (1)$$

onde:

$$f_1(j, k) = \begin{cases} 1, & \text{se } G_{n_i} = GE_{i_{jk}} \\ 0, & \text{se } G_{n_i} \neq GE_{i_{jk}} \end{cases} \quad (2)$$

2. TPO de um gesto G_{n_i}

Somatório de cada gesto G_{n_i} contabilizado apenas uma vez por voluntário na função F_i

$TPO_{G_{n_i}} = \frac{\text{Somatório de todos os gestos diferentes realizados por todos os voluntários na função } F_i}{\text{Somatório de todos os gestos diferentes realizados por todos os voluntários na função } F_i}$

$$TPO_{G_{n_i}} = \frac{\sum_{x=1}^{X_i} \sum_{t=1}^{T_{ix}} f_2(x, t)}{N_i} \quad (3)$$

onde:

$$f_2(j, k) = \begin{cases} 1, & \text{se } G_{n_i} = GDE_{ij_k} \\ 0, & \text{se } G_{n_i} \neq GDE_{ij_k} \end{cases} \quad (4)$$

3. TOV de um gesto G_{n_i}

Somatório de cada gesto G_{n_i} contabilizado apenas uma vez por voluntário na função F_i

$TOV_{G_{n_i}} = \frac{\text{Quantidade de todos os voluntários (experimentos) que realizaram os gestos para a função } F_i}{\text{Quantidade de todos os voluntários (experimentos) que realizaram os gestos para a função } F_i}$

$$TOV_{G_{n_i}} = \frac{\sum_{x=1}^{X_i} \sum_{t=1}^{T_{ix}} f_2(x, t)}{X_i} \quad (5)$$

4. TOT de um gesto G_{n_i}

Somatório de cada gesto G_{n_i} realizado na função F_i multiplicado pela diferença entre $TE_{ij_{Y_i}}$ e TE_{ij_k} (“tempo relativo” de G_{n_i})

$TOT_{G_{n_i}} = \frac{\text{Somatório de todos os tempos relativos dos } X_i's \text{ experimentos realizados para } F_i}{\text{Somatório de todos os tempos relativos dos } X_i's \text{ experimentos realizados para } F_i}$

$$TOT_{G_{n_i}} = \frac{\sum_{x=1}^{X_i} \sum_{z=1}^{Z_{ix}} f_3(x, z)}{\sum_{x=1}^{X_i} \sum_{z=1}^{Z_{ix}} TE_{ix_z}} \quad (6)$$

onde:

$$f_3(j, k) = \begin{cases} \left(TE_{ij_{Z_{ij}}} - TE_{ij_k} + 1 \right), & \text{se } G_{n_i} = GE_{ij_k} \\ 0, & \text{se } G_{n_i} \neq GE_{ij_k} \end{cases} \quad (7)$$

O fato de fazer $(TE_{i_j Z_{i_j}} - TE_{i_j k} + 1)$ é para que os gestos realizados primeiro tenham um maior peso, uma vez que $TE_{i_j Z_{i_j}}$ corresponde ao tempo de realização do gesto (Z_{i_j}) , que representa o último gesto realizado pelo voluntário E_{i_j} .

4.4 Escolha do vocabulário

O vocabulário de gestos é composto pelos gestos que melhor representam cada uma das funções. Assim, após analisar os gestos e calcular as taxas de ocorrência para cada um deles em cada uma das funções, deve-se determinar quais são esses gestos, sendo necessário assim, seguir seis etapas para obtê-los:

Etapla 1: Criar uma tabela para cada uma das taxas de ocorrência anteriormente calculadas. Tal tabela deve conter: o identificador do gesto, taxa de ocorrência correspondente e um peso, o qual será atribuído na etapa 3.

Etapla 2: Ordenar de maneira decrescente cada uma das tabelas, utilizando como parâmetro de ordenação a taxa de ocorrência de cada gesto.

Etapla 3: Considerando o modelo acima, onde para uma dada função F_i existam N_i gestos distintos, e que o gesto G_{n_i} corresponde ao gesto de posição n_i , sendo $1 \leq n_i \leq N_i$, deve-se atribuir um peso para o gesto G_{n_i} (denominado $P_{G_{n_i}}$), de maneira a fazer com que os gestos que estão nas primeiras posições de cada uma das quatro tabelas das taxas de ocorrência (TO), tenham uma maior chance de compor o vocabulário. Assim, o cálculo deste peso pode ser realizado da seguinte maneira:

$$P_{G_{n_i}} = \begin{cases} (N_i - n_i + 1) \frac{1}{N_i}, & \text{se } n_i = 1 \text{ ou } TO_{G_{n_i}} \neq TO_{G_{n_i-1}}, \forall n_i > 1 \\ P_{G_{n_i-1}}, & \text{se } TO_{G_{n_i}} = TO_{G_{n_i-1}}, \forall n_i > 1 \end{cases} \quad (8)$$

Etapla 4: Após atribuir os pesos para cada um dos gestos das quatro tabelas, os quais pertencem a um intervalo entre $(0,1]$, cria-se uma nova tabela contendo o identificador de cada um dos gestos e a média aritmética dos quatro pesos calculados na etapa anterior para cada gesto, que será denominado Peso Global (PG). O resultado de PG, o qual também pertence a um intervalo entre $(0,1]$, corresponderá ao nível de representatividade que cada um dos gestos possui perante os outros, sendo este, o único índice que deverá ser levado em consideração para a escolha do vocabulário.

Etapa 5: Após repetir a etapa anterior para todos os gestos, deve-se ordenar de maneira decrescente a tabela resultante, usando como parâmetro de ordenação o Peso Global.

Etapa 6: Por fim, basta escolher os gestos com melhores colocações de cada função para compor o vocabulário. É necessário notar que os gestos com mesmo PG têm o mesmo nível de representatividade, e por isso devem representar a mesma função no vocabulário.

No próximo capítulo esta metodologia será empregada na prática, a fim de demonstrar sua utilização de forma concreta.

5 Aplicação da Metodologia Proposta

Para poder validar a metodologia proposta, este capítulo apresenta como cada etapa foi executada, quais os resultados obtidos e, por meio da comparação com a metodologia de Nielsen, discute a qualidade do vocabulário obtido. Com a aplicação da metodologia proposta (a qual neste capítulo será chamada de metodologia 1) foi concebido um vocabulário de gestos intuitivos para interação entre pessoas e robôs móveis. Para isso, cada uma das etapas descritas na metodologia (definição das funções, captura dos dados gestuais, escolha do vocabulário e análise dos gestos) foi executada como se segue.

5.1 Etapa 1. Definição das Funções

Como este capítulo se propõe apenas a validar a metodologia, e não a resolver um problema específico, buscou-se na literatura quais as principais funções que os trabalhos sobre HRI abordavam ([KORTENKAMP; HUBER; BONASSO, 1996](#); [WALDHERR et al., 1998](#); [HASANUZZAMAN et al., 2005](#); [YANG; PARK; LEE, 2007](#); [BARATTINI; MORAND; ROBERTSON, 2012](#); [MEENA; JOKINEN; WILCOCK, 2012](#); [FUJII; LEE; OKAMOTO, 2014](#)). Desta maneira, foram selecionadas 7 funções: *abortar*, *apontar*, *atenção*, *mais devagar*, *mais rápido*, *ok* e *siga-me*. Estas são funções comuns para robôs móveis e podem ser entendidas como comandos que a interface pode passar para o robô, por isso o vocabulário aqui proposto pode ser utilizado por futuros trabalhos que desejem utilizar algumas delas em seus estudos.

5.2 Etapa 2. Captura dos dados gestuais

5.2.1 Definindo os parâmetros iniciais

Segundo [Jakob \(2000\)](#), são suficientes 5 experimentos para a realização de um teste de usabilidade com uma função. Apesar disso, neste trabalho foram considerados 12 experimentos para determinar os gestos que mais intuitivamente representassem cada uma das 7 funções escolhidas, perfazendo assim um total de 84 experimentos. Isto porque considera-se que quanto mais experimentos forem realizados, mais representativos serão os gestos capturados. No entanto, antes de realizá-los era necessário saber o seu tempo máximo de duração, levando em consideração que dentro desse tempo a frustração do voluntário

deveria ser alcançada. Então, foram realizados experimentos piloto com 7 participantes do Grupo de Pesquisa em Robótica da UFS (GPR-UFS), nos quais se observou que, em média, nos primeiros 18 segundos de experimento os voluntários já se frustravam por não conseguirem interagir com o robô, assim, para aumentar a chance de sucesso com a frustração e ao mesmo tempo minimizar o custo de armazenamento dos dados capturados, foi estipulado um tempo mínimo de 18 segundos e máximo de 30 segundos para a realização de cada experimento.

5.2.2 Escolha do sensor

Como os experimentos consistem na captura de dados de imagens a fim de serem analisadas posteriormente, em geral, esse procedimento poderia ser realizado por meio de uma câmera RGB (do inglês *Red, Green, Blue*) convencional. No entanto, como este trabalho também objetiva, em segundo plano, reconhecer os gestos escolhidos para compor o vocabulário, a escolha do sensor foi uma decisão muito importante, pois evitou dispendiosos tratamentos de imagens, uma vez que o sensor escolhido foi o *Microsoft Kinect 360* (ou apenas *Kinect*), mostrado na Figura 16.



Figura 16 – *Microsoft Kinect Sensor 360* (MICROSOFT, 2015)

A escolha do *Kinect* foi influenciada sobretudo por duas características principais do sensor: capacidade de fornecimento de múltiplos dados (imagem RGB, imagem de profundidade e o esqueleto virtual do indivíduo) e o baixo custo de aquisição.

Alguns trabalhos, como o de Pereira (2012), reconhece e identifica indivíduos por meio da junção de vários sensores. Este tipo de abordagem é eficaz, no entanto a reprodução de experimentos desse tipo pode demandar um investimento considerável, uma vez que alguns sensores adquiridos separadamente costumam ter um valor elevado. Isto, além de complicar a reprodução dos experimentos, também dificulta a sua aplicação em problemas reais. Por outro lado, o *Kinect*, além de ser considerado um sensor de baixo custo (SANTOS; NUNES, 2015), pode reconhecer vários indivíduos presentes na cena, identificá-los e ainda

oferecer dados que podem agregar mais informações para o sistema reconhecedor. Dentre esses dados, os principais são: imagem RGB, imagem de profundidade do ambiente e o esqueleto virtual do indivíduo, sendo este último descrito por meio de 21 pontos em 3 dimensões que representam suas principais articulações.

Outra característica desse sensor é a facilidade de implementação de códigos (NETO et al., 2013). Isto deve-se ao fato de que muitos projetos já foram construídos utilizando o *Kinect* como principal sensor (SANTOS; NUNES, 2015) e também pelo fato de que muitos outros trabalhos sobre HRI baseados em gestos já o utilizam como sensor para a aquisição de dados (DRESZER et al., 2011; MASTER; STOCKHOLM, 2011, 2011; KATALINIC, 2011), portanto existe uma facilidade enorme de se encontrar documentação sobre o mesmo.

Desta maneira, o *Kinect* foi o sensor escolhido para a realização da captura dos dados gestuais. Assim, é possível aproveitar o seu mapeamento de dados RGB-D e o esqueleto virtual, o que acaba facilitando o reconhecimento dos gestos. Além disso, será possível contribuir com a comunidade científica por meio da disponibilização de todos esses dados no sítio web www.gprufs.org, permitindo que outros pesquisadores que buscam iniciar pesquisas em HRI baseada em gestos, possam realizar seus estudos sem a necessidade de despender tempo e recursos para colher dados gestuais, sejam eles de imagens, profundidade ou de esqueleto virtual.

5.2.3 Os Voluntários

Os experimentos foram realizados com estudantes de 3 universidades distintas, sendo os mesmos, em sua maioria, do sexo masculino e da área de engenharias e computação. Cada voluntário assinou um termo de consentimento, no qual informava estar ou não de acordo com o aparecimento de suas imagens em possíveis divulgações de resultados obtidos através dos estudos realizados. Além deste termo, cada voluntário preencheu um formulário que possibilitou a análise do perfil de cada um (ambos os documentos podem ser visualizados no Apêndice F). Todos os dados catalogados podem ser vistos na Figura 17.

5.2.4 O Ambiente Utilizado

Para fazer com que cada voluntário tenha a impressão de estar realmente interagindo com o robô, o ambiente deve ser escolhido com muita cautela, pois uma má escolha pode influenciar nas decisões de cada voluntário. Outro aspecto que influenciou na escolha do ambiente foi o uso do *Kinect* como o sensor de captura de dados. Como este tem uma limitação quanto à distância para o usuário (intervalo entre 0,8 e 4 metros), para poder capturar os dados com melhor qualidade, o ambiente deve ser escolhido de maneira a garantir o mínimo de espaço necessário. Assim, foi estabelecido que na realização de cada experimento o voluntário ficaria a uma distância aproximada de 3,8 metros do *Kinect*, que, por meio de um suporte, ficava a uma altura de 1,1 metro do solo. O robô iniciava as tarefas a uma distância aproximada de 80 centímetros à frente do *Kinect* e a 3 metros de distância do voluntário, e se deslocava aproximadamente 2,5 metros até que o experimento acabasse, assim o voluntário era instruído a não avançar na direção do robô.

Devido à necessidade de ir ao encontro dos voluntários para realizar os experimentos, os mesmos foram improvisados em salas de aula ou laboratórios, onde o espaço necessário era consideravelmente pequeno. Por esse motivo, os locais tiveram de ser adaptados (apenas mudando alguns móveis de lugar) para oferecer ao menos o espaço mínimo (especificado nas medidas descritas previamente) que permitisse a interação entre o voluntário e o robô (Figura 18).

Sexo:		Faixa de Idade:		Área do Curso:		Trabalha?		Qual a frequência de uso de tecnologias?		Destro?		Já utilizou gestos para interagir com alguma tecnologia?		Joga vídeo Games ?	
Masculino	80,60%	18-19	34,33%	Exatas e da terra	5,97%	Sim	58,21%	Sempre	88,06%	Sim	83,58%	Sim	74,63%	Sempre	41,79%
Feminino	13,43%	20-29	46,27%	Humanas	0,00%	Não	35,82%	Às vezes	5,97%	Não	5,97%	Não	19,40%	Às vezes	40,30%
N/D	0,00%	30-39	10,45%	Sociais	0,00%	N.R	5,97%	Quase nunca	0,00%	N.R	10,45%	N.R	5,97%	Quase nunca	13,43%
N.R	5,97%	40-49	4,48%	Saúde	0,00%			N.R	5,97%					N.R	4,48%
		50-59	0,00%	Agrárias	0,00%										
		>=60	0,00%	Linguística, letras e artes	0,00%										
		N.R	4,48%	Biológicas	0,00%										
				Engenharias e Computação	88,06%										
				N.R	5,97%										

Figura 17 – Dados dos voluntários que realizaram os experimentos. Legenda: N/D - Não Definido, N.R - Não Responderam.

5.2.5 Os Experimentos

Em cada experimento o voluntário era convidado a interagir com um robô móvel do tipo *Pioneer 3-DX*. Para cada experimento o robô realizava uma tarefa predeterminada e o voluntário era instruído a interagir com o robô de modo a fazê-lo executar uma das 7 funções escolhidas. Para cada voluntário era dito que o robô deveria executar uma função e que esta era representada por um único gesto predefinido. Assim, o voluntário teria que realizar gestos até que “acertasse” qual era aquele que o robô reconhecia como o gesto correspondente à função. No entanto, por estar sendo controlado remotamente (mecanismo

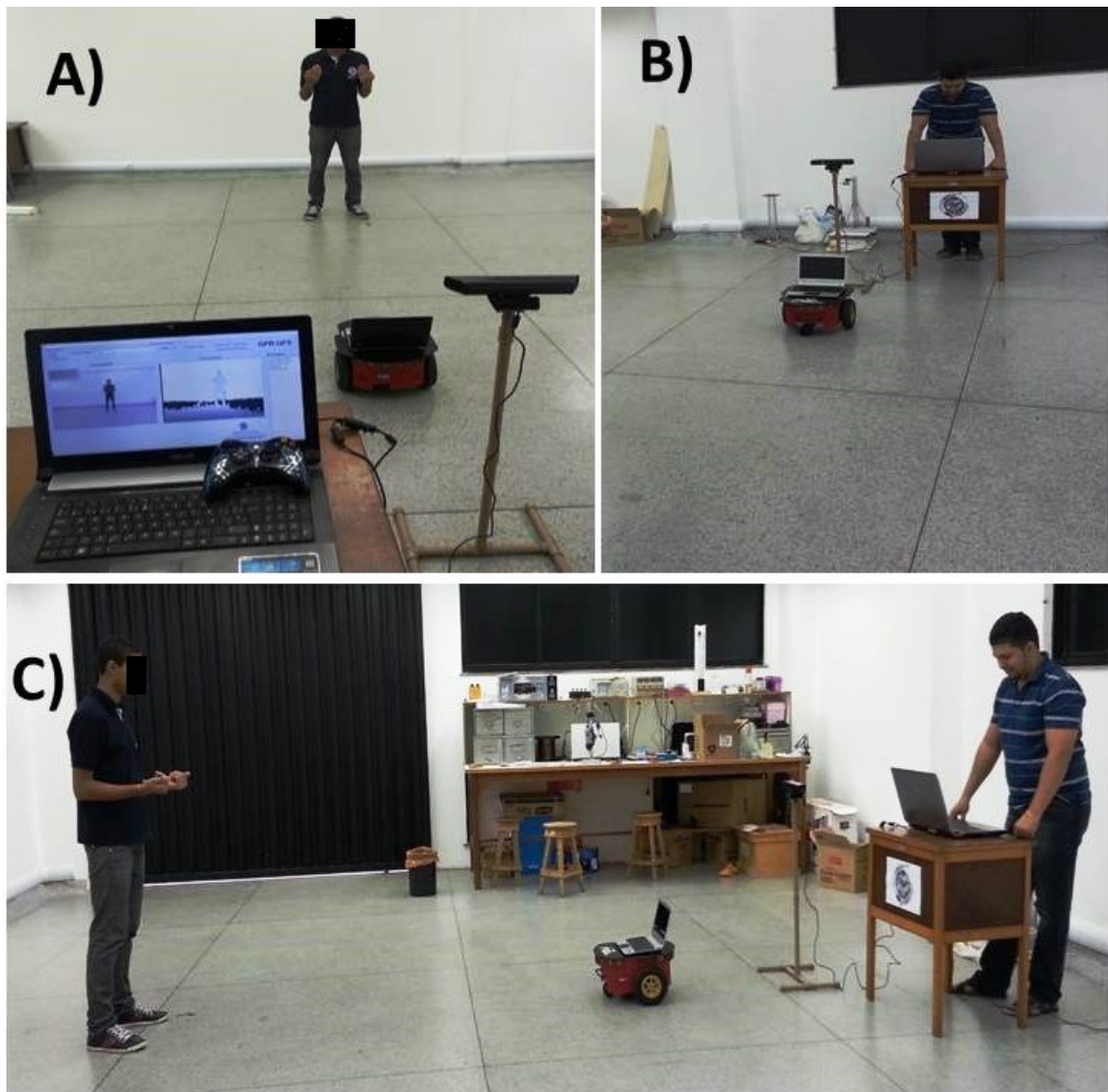


Figura 18 – Ambiente utilizado nos experimentos realizados no laboratório do GPR-UFS. A) visão do aplicativo do experimento. B) visão do voluntário. C) visão de um observador externo

definido pela abordagem do *Wizard of Oz*), o robô nunca respondia de maneira satisfatória aos comandos gestuais passados, o que levava o voluntário a realizar intuitivamente todos os gestos que lhe pareciam representar aquela função. Desta maneira, chegava o momento em que o voluntário se frustrava por não conseguir interagir com o robô e então o objetivo do experimento era alcançado. A frustração podia ser percebida no momento em que o voluntário questionava o mau funcionamento da interface ou então demonstrava não ter mais ideia de qual gesto realizar.

Antes de começar o experimento, para demonstrar ao voluntário a capacidade do robô de reagir a gestos, o voluntário era informado de que deveria realizar um determinado gesto, demonstrado pelo pesquisador, que indicaria ao robô que o mesmo deveria retroceder um pouco. O pesquisador então informava ao voluntário que estava habilitando a interface de comunicação baseada em gestos e fazia sinal para o voluntário de que o mesmo já podia realizar o gesto para o robô retroceder. Quando o pesquisador observava o gesto sendo realizado pelo voluntário, comandava o robô para que o mesmo executasse a função de retroceder, dando a certeza ao voluntário de que o robô poderia mesmo responder a gestos, desde que ele de fato fosse capaz de realizar o gesto correto para a função em questão.

Sendo assim, como indicado na metodologia proposta, cada voluntário realizou apenas um experimento o qual corresponde à apenas uma função. Para cada uma das 7 funções definidas os voluntários foram instruídos da seguinte maneira:

1. **Abortar:** O objetivo dessa função é fazer com que o robô pare de executar uma tarefa independente do ponto de execução em que ela esteja. Para isso, considerando que a tarefa designada ao robô foi a de ir na direção do voluntário, este teria que realizar gestos que o robô entendesse como um comando para ele parar de se mover. No entanto, o robô continuava se movendo na direção do voluntário até o término do experimento, que se dava quando o voluntário se frustrava ou depois de decorridos 25 segundos do início do experimento.
2. **Atenção:** O objetivo desta função é fazer com que o robô perceba que alguém está chamando sua atenção. Assim, enquanto o robô ficava parado, o voluntário era instruído a realizar gestos até que o robô percebesse sua presença e se movesse. No entanto, o robô permanecia parado até o término do experimento, que se dava quando o voluntário se frustrava ou depois de decorridos 30 segundos do início do experimento.

3. **Apontar:** Essa função objetiva fazer o robô dirigir-se até um ponto específico no ambiente. Para isso, enquanto o robô permanecia parado, o voluntário era instruído a apontar para um ponto qualquer no ambiente o qual ele desejasse que o robô se dirigisse. No entanto, o robô permanecia parado até o término do experimento, que se dava pela frustração do voluntário ou depois de decorridos 25 segundos do início do experimento.
4. **Mais devagar:** O objetivo desta função é fazer o robô diminuir sua velocidade. Assim, o voluntário foi instruído a realizar gestos que sinalizassem para o robô que ele deveria diminuir a sua velocidade atual. No entanto isso nunca ocorria, uma vez que o robô andava a uma velocidade linear fixa de 12 cm/s e só parava quando o experimento acabava, o que se dava quando o voluntário se frustrava ou depois de decorridos 22 segundos do início do experimento.
5. **Mais rápido:** Esta função é contrária à anterior, pois o objetivo é fazer o robô aumentar sua velocidade. Para isso, o voluntário era instruído a realizar gestos que sinalizassem para o robô que ele deveria aumentar a sua velocidade atual. No entanto isso nunca ocorria, já que o robô se dirigia na direção do voluntário a uma velocidade fixa de 5 cm/s e só parava com o término do experimento, que se dava pela frustração do voluntário ou depois de decorridos 25 segundos do início do experimento.
6. **Ok:** O objetivo desta função é fazer com o robô receba do voluntário um retorno positivo correspondente à qualidade com que uma determinada tarefa foi executada. Neste caso, enquanto o robô se dirigia na direção do voluntário, este deveria realizar gestos sinalizando que a tarefa (dirigir-se na sua direção) fora realizada com sucesso. Para isso, enquanto o robô se dirigia na direção do voluntário, este foi instruído a realizar os gestos de *ok* até que o robô reconhecesse e parasse. No entanto o robô nunca demonstrava ter recebido esta informação, pois ele continuava se movendo na direção do voluntário até o término do experimento, que se dava quando o voluntário se frustrava ou depois de decorridos 25 segundos do início do experimento.
7. **Siga-me:** O objetivo dessa função é fazer o robô seguir o voluntário. Para isso, enquanto o robô ficava parado, o voluntário devia realizar gestos que sinalizassem para o robô que este deveria segui-lo. No entanto, o robô nunca se movimentava e o experimento acabava pela frustração do voluntário ou depois de decorridos 30 segundos do início do experimento.

Para controlar o robô, sinais de um joystick eram enviados remotamente para um aplicativo que estava sendo executado em um computador a bordo do mesmo. Desta maneira, apesar de parecer para o voluntário que o robô era autônomo e interagia por meio de gestos, ele apenas executava tarefas predefinidas.

Já para fazer o gerenciamento e captura dos dados de vídeo, foi utilizado outro computador, o qual, por meio de uma aplicação JAVA, recebia os dados capturados pelo *Kinect*, os tratava e os armazenava no formato de arquivos. Assim, após o término de cada experimento, tinha-se armazenados os dados de imagem, profundidade e esqueleto do voluntário que o realizou. Tais dados foram utilizados na análise e escolha do vocabulário, no reconhecimento dos gestos escolhidos e poderão ainda servir como base de dados para estudos sobre reconhecedores de gestos, os quais, mesmo não fazendo parte do escopo deste trabalho, podem ser de grande contribuição para outros estudos que estejam focados nos aspectos da máquina e não no usuário.

5.3 Etapa 3. Análise dos gestos

Cada experimento realizado foi armazenado de maneira a permitir a reprodução dos vídeos capturados. Assim, foi possível analisar cada um dos vídeos e criar as tabelas propostas na metodologia. A construção passo a passo de tais tabelas pode ser encontrada no Apêndice A, onde uma explicação detalhada deste processo de aplicação da metodologia proposta é apresentada.

Na análise realizada foram encontrados 97 gestos diferentes (que podem ser visualizados na Tabela 5 do Apêndice B), os quais foram identificados e descritos de maneira cautelosa, visando minimizar o efeito da subjetividade inerente ao processo no resultado da análise. Nesse processo, cada segundo de vídeo levou em média 30 segundos para ser analisado. Sendo assim, o total de 36 minutos de vídeos capturados nos 84 experimentos levaram em torno de 18 horas para serem analisados. É importante ressaltar que o processo de segmentação e descrição dos gestos foi realizado de forma manual. Todos os dados da análise foram salvos em um banco de dados, o que permitiu a utilização de instruções SQL para cálculo das taxas de ocorrência e do valor do peso geral de cada gesto em cada uma das funções. Entretanto, o uso de ferramentas de banco de dados não é uma exigência da metodologia proposta, podendo o usuário proceder à análise dos dados da forma que lhe pareça mais conveniente.

5.4 Etapa 4. Escolha do vocabulário

Após análise e cálculo do peso geral (PG) de cada gesto, foi possível ter, para cada uma das funções, uma lista de gestos que a representam, ordenados de maneira decrescente por nível de intuitividade (PG). Dessa maneira, a Figura 19 ilustra o vocabulário obtido pela escolha dos gestos com o maior PG em cada função.

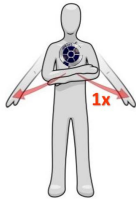
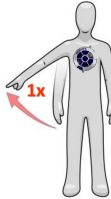
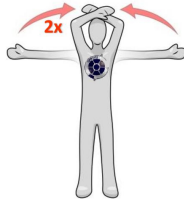


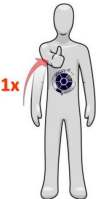
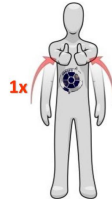

Abortar	Apontar	Atenção	Mais Devagar
			
Mais Rápido	Ok		Siga-me
			

Figura 19 – Vocabulário 1: Vocabulário obtido com a aplicação da metodologia 1

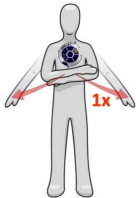
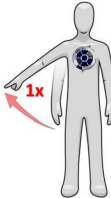



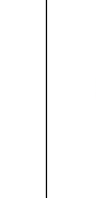
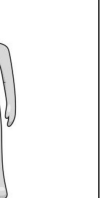

Abortar	Apontar	Atenção	Mais Devagar
			
Mais Rápido	Ok		Siga-me
			

Figura 20 – Vocabulário 2: Vocabulário obtido com a aplicação da metodologia 2

Como pode ser visto na Figura 19, a função “Ok” está sendo representada por dois gestos. Isto se deve ao fato dos dois gestos terem o mesmo PG e por isso, para este trabalho, ambos possuem o mesmo nível de intuitividade.

Para obter mais detalhes sobre a aplicação da metodologia, o Apêndice [A](#) descreve em detalhes todo o roteiro utilizado e os apêndices [B](#), [C](#), [D](#) e [E](#) descrevem todas as tabelas resultantes do processo de análise.

5.5 Aplicação da metodologia de Nielsen (metodologia 2)

Após a obtenção do vocabulário representado na Figura [19](#), o qual foi resultado da aplicação da metodologia 1, foi obtido um outro vocabulário, aplicando agora a metodologia 2. Como mencionado anteriormente, a metodologia 2 já é bastante utilizada em trabalhos de HCI, além de ser a única que é descrita de forma a permitir a sua reprodução, e por isso foi escolhida para ser comparada com a metodologia proposta. Nesse segundo processo foram utilizados os mesmos vídeos resultantes da aplicação da metodologia 1, no entanto, foi considerado apenas o primeiro gesto realizado por cada voluntário, conforme proposto pela metodologia 2. Assim, após calcular a frequência dos primeiros gestos realizados por cada voluntário, foi possível obter outro vocabulário, o qual está ilustrado na Figura [20](#).

5.6 Análise comparativa entre as metodologias 1 e 2

Analisando os vocabulários 1 e 2, obtidos respectivamente pelas duas metodologias utilizadas, mesmo havendo algumas diferenças entre eles, como por exemplo a diferença entre os gestos da função “mais devagar” e a função “ok”, que no vocabulário 2 é representada por apenas um gesto, os vocabulários são muito semelhantes, o que é um indício de que a metodologia 1 pode trazer resultados pelo menos tão bons quanto a metodologia 2, que já é amplamente aceita em HCI.

Diante desses resultados, existem algumas considerações a serem feitas:

1. por questão de cumprimento de prazos para a finalização do presente trabalho, os vocabulários não foram avaliados a fim de medir o grau de intuitividade de cada um;
2. a metodologia 1 encontra resultados por meio de experimentos um pouco mais demorados que a metodologia 2, no entanto, ao aplicá-la, o pesquisador além de obter uma maior variedade de gestos intuitivos, pode escolher mais de um gesto para cada função, o que aumenta a complexidade do vocabulário e possivelmente influencia diretamente no aumento da robustez do mesmo;

3. o peso de cada gesto é calculado por meio de 4 diferentes taxas de ocorrência, as quais medem as características de cada gesto em aspectos distintos (ocorrência geral, ocorrência por voluntário, ocorrência entre todos os voluntários e ocorrência por tempo), o que não é observado na metodologia 2, já que nesta os gestos que devem compor o vocabulário são escolhidos usando apenas a medida de frequência de aparecimento dos mesmos, fazendo com que outros aspectos importantes, como estes citados, sejam desconsiderados;
4. ao aplicar ambas metodologias, os experimentos para encontrar os gestos devem ser planejados de maneira cuidadosa, uma vez que se os mesmos forem executados de uma maneira equivocada, os resultados podem ser bem diferentes daquilo que se busca, que é um vocabulário intuitivo.
5. É importante destacar que tanto no caso da metodologia 1 como no caso da metodologia 2, os gestos que são realizados com um único braço, tanto podem ser feitos com o braço esquerdo como com o braço direito, apesar disto não estar claramente ilustrado nas figuras 19 e 20.

Como pode ser visto na Tabela 11 do Apêndice E, a metodologia 1 obtém um *ranking* dos gestos que são mais prováveis de serem intuitivos para cada função. Dessa maneira, ao oferecer para o pesquisador a possibilidade de ter uma função representada por mais de um gesto (chamado de vocabulário complexo por (WACHS, 2006)), possivelmente o usuário terá mais chances de passar para o robô qual o comando ele quer que este execute. Considerando, hipoteticamente, que o reconhecedor de gestos tenha taxa de acerto de 100%, caso se tenha mais de um gesto intuitivo representando uma mesma função, a probabilidade do usuário realizar um gesto que o robô reconhece é bem maior do que se houvesse apenas um gesto representando cada função. Assim, é possível alcançar uma maior eficácia na interação e consequentemente uma melhor comunicação entre pessoas e robôs.

Portanto, pode-se perceber que a metodologia 1, mesmo sendo nova, além de poder levar a resultados semelhantes aos da metodologia 2, já utilizada e aceita pela HCI, ainda possui diferenciais, como a possibilidade de obter vocabulários mais complexos, que podem levar a vocabulários de gestos mais intuitivos e mais robustos. Além disso, percebe-se ainda que esse tipo de vocabulário pode ser um ótimo recurso para romper a barreira

existente entre as pessoas comuns e os robôs, e fazer com que a robótica adentre com mais efetividade na vida cotidiana das pessoas.

5.7 Classificação dos gestos dos vocabulários segundo a psicolinguística

Como pôde ser visto no Capítulo 2, os gestos podem ser classificados como icônicos, metafóricos, dêiticos, batidas e emblemáticos.

Dessa maneira, é possível classificar os gestos que compõem respectivamente os vocabulários 1 e 2 da seguinte maneira:

5.7.1 Classificação dos gestos do vocabulário 1

1. **Abortar:** O gesto que representa esta função não pode ser metafórico, pois não está descrevendo um objeto e um evento ao mesmo tempo. Não pode ser icônico, pois, apesar de estar representando uma ação, a mesma não remete a um aspecto imagístico que possa descrever a função Abortar. Não pode ser dêitico, pois não está apontando para um ponto específico do ambiente. Também não se trata de um gesto de batida. Sendo assim, como ele está sendo utilizado para passar um comando para o robô, esse gesto é classificado como emblemático.
2. **Apontar:** Por definição, o gesto que está representando esta função é um gesto dêitico.
3. **Atenção:** O gesto que representa essa função é um gesto emblemático. Os argumentos que justificam essa afirmação são os mesmo utilizados para classificar o gesto que representa a função Abortar.
4. **Mais Devagar:** Esse gesto descreve inicialmente uma ação sem que haja um objeto envolvido na mesma. Por esse motivo, ele não pode ser metafórico. Como o gesto não é de apontar para um ponto específico do ambiente, não pode ser classificado como dêitico. Também não é um gesto de batida. Ele poderia ser emblemático porque passa um comando para o robô. No entanto, ele executa uma ação que remete a um significado imagístico de recuo. Então, como o robô está se aproximando com uma velocidade relativamente alta, o indivíduo demonstra por meio desse gesto o desejo

de que essa velocidade diminua. Dessa maneira este gesto é entendido ser um gesto icônico.

5. **Mais Rápido:** O gesto que representa essa função também é icônico. A ação que ele representa remete a um significado imagístico de que o indivíduo quer que ele continue se aproximando, e que por haver repetições rápidas, o indivíduo demonstra o desejo de que essa aproximação seja feita com uma maior velocidade. Os argumentos que não o fazem ser classificado como outro tipo de gesto são os mesmos utilizados para a função “Mais Devagar”.
6. **Ok:** Como pode ser visto em (MCNEILL, 1992), os dois gestos que representam essa função são classificados como emblemáticos.
7. **Siga-me:** Esse gesto também pode ser considerado icônico, pois, uma vez que o indivíduo descreve uma trajetória que sai da sua frente e vai até o seu corpo, ele indica o desejo de fazer com que o robô se aproxime dele.

5.7.2 Classificação dos gestos do vocabulário 2

Como os dois vocabulários são praticamente iguais, os mesmos argumentos utilizados para classificar os gestos que representam as funções do vocabulário 1, podem ser utilizados para classificar os gestos que representam as funções do vocabulário 2. Dessa maneira, os gestos do vocabulário 2 podem ser classificados como:

1. **Abortar:** Emblemático.
2. **Apontar:** Dêitico.
3. **Atenção:** Emblemático.
4. **Mais Devagar:** Icônico.
5. **mais Rápido:** Icônico.
6. **Ok:** Emblemático.
7. **Siga-me:** Icônico.

5.7.3 Considerações sobre a classificação dos gestos

Mesmo os experimentos sendo executados por meio de gestos, em nenhum momento foi imposta para os voluntário uma restrição quanto à impossibilidade do uso da fala

durante o processo de experimentação. Dessa maneira, mesmo que a fala não tenha sido levada em consideração na análise dos gestos obtidos, esta pode ter estado presente na interação. Dessa forma, alguns gestos que dependem diretamente do fluxo de fala como os dêiticos, emblemáticos e os de batidas, podem se fazer presentes nos vocabulários obtidos como resultados desses experimentos.

6 Reconhecimento dos Gestos Obtidos no Vocabulário

De acordo com [Mitra e Acharya \(2007\)](#), o reconhecimento de gestos é um processo que pode ser entendido como o reconhecimento significativo das expressões passadas pelos movimentos realizados por um ser humano. Esses movimentos podem ser realizados pelos braços, pernas, mãos, cabeça, face e/ou todo o corpo. Dessa maneira, para que as máquinas possam interagir com os seres humanos por meio de gestos, é necessário que elas adquiram a capacidade de reconhecê-los. Isto pode ser alcançado por meio do reconhecimento de alguns padrões de movimento.

Segundo [McNeill \(1992\)](#), a maioria dos gestos são realizados pelos movimentos dos braços e das mãos. Sendo assim, grande parte das pesquisas está focada no reconhecimento dos gestos realizados pelos membros superiores. Estes gestos podem ser estáticos ou dinâmicos, e a depender dessa característica, o modelo de aprendizagem pode ser bem diferente.

Como a contribuição principal deste trabalho não se dirige ao aspecto do reconhecimento dos gestos obtidos no vocabulário, a utilização de um modelo que permita reconhecê-los já é suficiente. Assim, por motivos de facilidade de implementação, resolveu-se utilizar o DTW (do inglês - *Dynamic Time Warping*) como método de cálculo da distância entre os gestos e o KNN (do inglês *K-nearest neighbors*) como método de classificação.

6.1 Conhecendo o DTW

O DTW é um método utilizado para obter a distância entre duas séries temporais. Com ele, por meio dessa distância, é possível determinar o quão semelhantes são duas séries, mesmo que elas possuam diferentes distorções durante sua evolução no tempo. Assim, é possível verificar se um conjunto de dados que evolui no tempo é semelhante a outro conjunto de dados que contém as mesmas características.

Um DTW, em linha gerais, funciona da seguinte maneira:

Sejam duas séries temporais, A e B , de tamanho n e m respectivamente, onde:

$$A = a_1, a_2, \dots, a_n \quad (9)$$

$$B = b_1, b_2, \dots, b_m \quad (10)$$

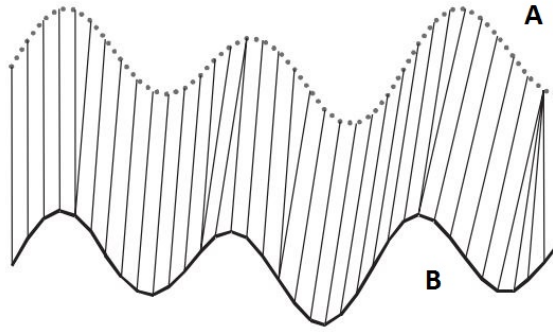


Figura 21 – Exemplo do alinhamento de duas séries quaisquer A e B

Assim, para alinhar as séries A e B , utiliza-se uma matriz W de dimensões $n \times m$ onde cada elemento (i, j) da matriz W representa a distância $d(a_i, b_j)$ entre os dois pontos a_i e b_j . Cada elemento (i, j) também corresponde ao alinhamento entre a_i e b_j , como pode ser visto na Figura 21. Sendo assim, após alinhar todos os pontos a_i e b_j pode-se obter um caminho P com K pontos que representam a menor distância entre as duas séries. Logo, em um caminho P , cada k 'ésimo ponto pode ser representado como $P_k = (i, j)_k$, de modo que

$$P = p_1, p_2, \dots, p_k, \dots, p_K \quad \max(m, n) \leq K \leq m + n - 1. \quad (11)$$

No entanto, antes de comparar duas séries, algumas condições devem ser satisfeitas:

1. **Fronteiras:** $P_1 = (1, 1)$ e $P_K = (m, n)$, assim o caminho P inicia e termina nas extremidades da diagonal secundária da matriz W .
2. **Continuidade:** Considerando $p_k = (x, y)$ e $p_{k-1} = (x', y')$, deve-se garantir que $x - x' \leq 1$ e $y - y' \leq 1$. Assim, é possível restringir o caminho P a seguir sempre por pontos adjacentes.
3. **Monotonicidade:** Considerando $p_k = (x, y)$ e $p_{k-1} = (x', y')$, deve-se garantir que $x - x' \geq 0$ e $y - y' \geq 0$. Assim, é possível forçar o caminho P a ser monotonicamente espaçado no tempo.

Como pode ser percebido, na matriz W existem exponencialmente vários caminhos que podem satisfazer as três condições citadas acima. No entanto, o caminho P que realmente interessa é aquele com o menor custo de distorção da série no tempo. Dessa forma, por ter uma complexidade exponencial, o algoritmo responsável por encontrar tal caminho pode ser otimizado por meio da programação dinâmica, de maneira a avaliar sempre a próxima

recorrência, que define a distância acumulada $C(i, j)$ como a distância $d(i, j)$ encontrada por meio da soma da distância do elemento atual (i, j) e da distância mínima acumulada pelos elementos adjacentes $(i - 1, j - 1)$, $(i - 1, j)$ e $(i, j - 1)$, como pode ser visto na expressão 12.

$$C(i, j) = d(a_i, b_j) + \min\{C(i - 1, j - 1), C(i - 1, j), C(i, j - 1)\}. \quad (12)$$

No DTW, deve-se levar em consideração que normalmente se utiliza a distância euclidiana para poder calcular cada $d(i, j)$:

$$d(i, j) = \sqrt{(a_i - b_i)^2 + (a_j - b_j)^2}. \quad (13)$$

No entanto várias outras medidas de distância podem ser utilizados, como por exemplo a distância *Manhattan*, que pode ser calculada da seguinte maneira:

$$d(i, j) = \|a_i - b_i\| + \|a_j - b_j\|. \quad (14)$$

A implementação do DTW é muito simples e ao mesmo tempo oferece um bom potencial de classificação. Assim sendo, o Algoritmo 1 faz uso da distância euclidiana para ilustrar a simplicidade de implementação de um DTW.

Algoritmo 1: Algoritmo básico de um DTW

```

Data: Serie A, Serie B
Result: D //Custo global da associação entre A e B
1 N=tamanho(A) //tamanho da serie A;
2 M=tamanho(B) //tamanho da serie B ;
3 DTW=[n+1,m+1] //inicia a matriz de custos;
4 DTW(1,:)=0 //faz a primeira linha iniciar com zero;
5 DTW(:,1)=0 //faz a primeira coluna iniciar com zero;
6 for i=1  $\Rightarrow$  n do
7   for j=1  $\Rightarrow$  m do
8      $\quad dist = dist\_euclidiana(A(i), B(j));$ 
9      $\quad DTW(i, j) = dist + \min\{DTW(i-1, j-1), DTW(i-1, j), DTW(i, j-1)\};$ 
10  D=DTW(n, m) //retorno;
```

Um gesto pode ser representado como um vetor de pontos que evolui no tempo de maneira a formar seu movimento característico. Assim, um gesto A é formado por um conjunto de vetores a_i , onde i representa a ordem temporal de obtenção dos pontos do vetor a_i . Logo, para comparar dois gestos A e B utilizando DTW, deve-se tomar um dos gestos como modelo e aplicar o Algoritmo 1 para verificar a distorção entre eles. Como

visto no Algoritmo 1, o DTW entre duas séries é um número que representa a distorção entre eles. Assim, no reconhecimento de gestos, costuma-se tomar alguns gestos como modelo de comparação e calcula-se o DTW em um gesto a ser classificado e cada um dos modelos de comparação. Logo, o gesto pode ser classificado como o gesto modelo que resultou no menor valor de distorção.

6.2 Conhecendo o KNN

O KNN é um algoritmo de classificação que se baseia no cálculo de distâncias entre dados (COVER; HART, 1967). Para classificar um padrão X , o algoritmo KNN calcula os K -vizinhos mais próximos a X e classifica-o como sendo da classe que aparece com maior frequência dentre os seus K -vizinhos. É importante observar que quando a quantidade de padrões de cada classe é diferente, cada voto de cada um dos k -vizinhos, deve ser ponderado pelo inverso da quantidade de padrões que a classe a qual esse vizinho pertence possui.

Como pode ser visto no Algoritmo 2, assim como o DTW, o KNN também é um algoritmo simples de implementar. No entanto, essa simplicidade não interfere na sua capacidade de oferecer bons resultados de classificação.

Algoritmo 2: Algoritmo básico de um KNN

```

Data: Serie A, Serie B, k
1 //A=testes B=treinamento K=vizinhos;
Result: D //lista de classes
2 N=tamanho(A) //tamanho da serie A;
3 M=tamanho(B) //tamanho da serie B ;
4 Vetor D(N) //Vetor para armazenar as N classes;
5 for  $i=1 \Rightarrow n$  do
6   Vetor distancias(M) //vetor de distâncias de tamanho M;
7   for  $j=1 \Rightarrow m$  do
8      $dist = \text{calcula\_distancia}(A(i), B(j))$ 
9   ordena_crescente(distancia) //ordena em ordem decrescente o vetor de
    distâncias;
10  Vetor vizinhos(k) //vetor para armazenar as distâncias dos k-vizinhos;
11  vizinhos=pega_vizinhos_proximos(K) //armazena em vizinhos os k mais
    próximos;
12  classe  $c = \text{frequencia\_max}(vizinhos)$  //obtem a classe com maior frequência;
13   $D(i) = c$ 

```

6.3 Reconhecimento dos gestos

Como visto no Capítulo 5, o vocabulário encontrado com a aplicação da metodologia proposta é constituído de oito gestos que representam sete funções. Para demonstrar que estes gestos podem ser implantados em um robô real, existe a necessidade de demonstrar que os mesmos são factíveis de reconhecimento por parte da máquina.

Assim, sendo um gesto representado por um vetor de pontos que evolui no tempo, para determinar quais dados deveriam fazer desse vetor, foi necessário atentar para as seguintes observações:

- Os dados gestuais foram capturados pelo sensor *Kinect*. Assim, têm-se os dados de imagem, profundidade e do esqueleto virtual de cada voluntário, disponíveis para o reconhecimento dos gestos. Logo, seria possível utilizar as imagens e/ou os dados de profundidade para tal fim. Porém, seria necessária a aplicação de uma etapa de pré-processamento para melhorar a qualidade dos dados e extrair suas características mais significativas. Por outro lado, os dados de esqueleto virtual, além de serem ricos em informação, já estão prontos para serem utilizados, o que dispensa uma etapa extra de pré-processamento.
- Cada junta do esqueleto virtual (ponto de intercessão entre dois segmentos, Figura 22) é um ponto em 3D, onde:
 - A origem das coordenadas (x, y) está localizada no centro na imagem, enquanto a coordenada z , por representar a distância existente entre ponto e o *Kinect*, tem sua origem coincidindo com a posição do sensor. Assim, a depender da localização dos voluntários na cena, movimentos iguais podem ser representados por conjuntos de pontos bem diferentes. Portanto, para um reconhecimento mais efetivo, é imprescindível que os movimentos sejam invariantes à posição de cada voluntário, ou que pelo menos essa variação seja mínima. Por esse motivo, resolveu-se tomar o ponto do torso (identificado na Figura 22) de cada voluntário como a origem de um novo sistema de coordenadas. Desta maneira, independentemente de onde o voluntário esteja, o referencial variará o mínimo possível.
 - Os pontos centrais das mãos nem sempre representam movimentos diferenciáveis, pois, a depender de outras informações, como por exemplo a posição dos coto-

velos, os gestos podem ser diferentes. Assim, existe a necessidade de acrescentar outros dados nos vetores que representam os movimentos gestuais. Essas informações adicionais objetivam facilitar a distinção entre os gestos. Dessa maneira, cada gesto a ser reconhecido foi representado pela evolução no tempo de um vetor com 8 dimensões, com elementos dispostos na seguinte ordem:

1. Coordenada x da mão direita
2. Coordenada y da mão direita
3. Coordenada z da mão direita
4. Ângulo formado entre os pontos do ombro, cotovelo e da mão do braço direito
5. Coordenada x da mão esquerda
6. Coordenada y da mão esquerda
7. Coordenada z da mão esquerda
8. Ângulo formado entre os pontos do ombro, cotovelo e da mão do braço esquerdo

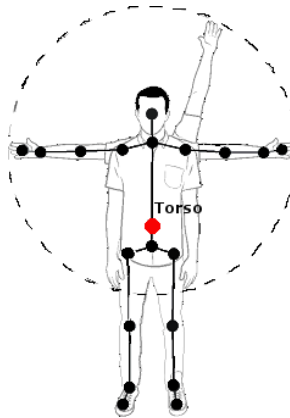


Figura 22 – Representação de um esqueleto virtual (CORP, 2016)

Logo, após representar os gestos em series de vetores, foram escolhidos os dois modelos (DTW e KNN) para poder classificá-los.

6.3.1 Treinamento, testes e resultados

No processo de classificação foi utilizado um conjunto de 179 gestos que se dividem, de forma não uniforme, entre os oito gestos diferentes que compõem o vocabulário obtido pela aplicação da metodologia proposta neste trabalho, conforme indicado na Tabela 2.

O processo de classificação se deu pela validação cruzada do tipo *leave-one-out* (DUDA; HART; STORK, 2012), onde cada um dos gestos foi comparado com os outros 178 gestos. Para isso, o DTW foi utilizado para calcular a distância entre os conjuntos de vetores que representam os gestos e o KNN foi utilizado para a classificação.

Para a determinação do K , número que representa os k -vizinhos mais próximos de um dado gesto, foi realizado um experimento onde todos os K 's, de 1 à 178, foram testados. Dessa maneira, a maior taxa de acerto foi obtida com K igual a 9.

Tabela 2 – Quantidade de gestos para cada função usados na classificação.

Função	Quantidade de gestos
Abortar	45
Apontar	24
Atenção	14
Mais Devagar	27
Mais Rápido	9
Ok (com apenas um braço)	16
OK (com os dois braços)	12
Siga-me	32

Considerando o pior caso, onde cada uma das sete classes (Tabela 2) poderia ser escolhida ao azar, a probabilidade de acerto é de aproximadamente 14,5%. No entanto, após o processo de classificação descrito acima, foi obtida uma taxa média de acertos de 77,5%. Apesar dessa taxa não ser a ideal para a implantação dos gestos em um ambiente interacional real, existem alguns aspectos a serem considerados que possivelmente influenciaram este resultado, a saber:

- O número de dados utilizados no reconhecimento foi pequeno, o que pode influenciar diretamente na variabilidade dos modelos utilizados. Isto, por consequência, pode aumentar a frequência de aparecimento de falsos-positivos.
- Os gestos não foram coletados em um ambiente controlado, assim os usuários não executaram gestos bem elaborados para melhorar a qualidade dos dados a serem

fornecidos ao reconhecedor. Dessa maneira, um mesmo gesto varia de usuário para usuário, o que acaba aumentando as chances de confusão entre eles.

- Alguns gestos se valem de uma determinada característica para expressarem seu significado. Esse é o caso dos gestos representantes da função “OK” do vocabulário, onde, além da trajetória, uma característica fundamental que o diferencia é a extensão dos dedos polegares. Já no gesto que representa a função “mais devagar”, a característica principal que o diferencia é a extensão de todos os dedos da mão. Dessa forma, caso não seja possível identificar tais características, o gesto possivelmente não irá ser reconhecido. Como o sensor utilizado na obtenção dos gestos foi o *Kinect 360*, e ele não possui a capacidade de identificar o dedos das mãos dos usuários, tornou-se difícil classificar alguns gestos de maneira correta apenas através das informações de suas trajetórias. Apenas a título de informação, o *Kinect 2* tem a capacidade de identificar a posição dos dedos das mãos.
- Alguns gestos são caracterizados também pelo número de repetições do movimento e pela sua velocidade de realização. Assim, como esses aspectos não foram tratados no momento do reconhecimento, alguns gestos, como os que representam as funções “Atenção” e “Mais Devagar”, por exemplo, acabaram sendo confundidos com outros.
- O modelo utilizado sugere reconhecer gestos realizados com os dois braços. Dessa maneira, como nos trabalhos sobre reconhecimento de gestos dinâmicos não foi encontrado um modelo que reconhecesse gestos realizados com as duas mãos, o mesmo deve ser estudado com mais atenção, pois, ele demonstra ser uma boa alternativa para o reconhecimento de gestos de corpo inteiro, o que, segundo [Yang, Park e Lee \(2007\)](#), ainda é um desafio a ser enfrentado para interação natural.

Para fornecer mais informações sobre o reconhecimento de gestos, a tabela 3 lista trabalhos que se dedicam a apresentar alguns modelos de reconhecimento de gestos.

Tabela 3 – Trabalhos que objetivam desenvolver reconhecedores de gestos

Trabalho	Tipo de gesto	Método Utilizado
(BETTIO et al., 2013)	Dinâmico	FSM + SVM
(BHUYAN, 2012)	Dinâmico	FSM
(BOBICK; SOCIETY, 1997)	Dinâmico	FSM

Trabalho	Tipo de gesto	Método Utilizado
(DAVIS; SHAH, 1994)	Dinâmico	FSM
(HONG et al., 2000)	Dinâmico	FSM + K-means
(VERMA; DEV, 2009)	Dinâmico	FSM + Fuzzy c-means
(YEASIN; CHAUDHURI, 2000)	Dinâmico	FSM
(WU; WANG; ZHANG, 2015)	Dinâmico	HMM + Fuzzy ANN
(BOWDEN et al., 2004)	Dinâmico	HMM + ICA
(ELMEZAIN; AL-HAMADI, 2009)	Dinâmico	HMM
(ELMEZAIN et al., 2008)	Dinâmico	HMM
(KIM, 1999)	Dinâmico	HMM
(KIM; JANG, 2015)	Dinâmico	HMM
(LIU et al., 2014)	Dinâmico	Múltiplos HMM's
(RASHID; AL-HAMADI; MICHAELIS, 2009)	Dinâmico	HMM + SVM
(STARNER; PENTLAND, 1996)	Dinâmico	HMM
(WANG et al., 2012)	Dinâmico	HMM
(YAMATO; OHYA; ISHII, 1992)	Dinâmico	HMM
(ZHANG et al., 2014)	Dinâmico	DDF + HMM
(KOTA et al., 2009)	Estático	PCA
(BHOKSE; KARWANKAR, 2015)	Estático	ANN
(LEE et al., 2004)	Dinâmico	Entropia
(GAUS; WONG, 2012)	Dinâmico	Filtro de Kalman
(TRAN et al., 2014)	Dinâmico	MHI + KDES + SVM
(AWARD; SUTHERLAND, 2006)	Estático	SVM
(BISWAS; BASU, 2011)	Estático	SVM
(CHEN; TSENG, 2007)	Estático	SVM
(ALON et al., 2007)	Dinâmico	DSTW
(STEFAN et al., 2008)	Dinâmico	DSTW

7 Conclusões e Trabalhos Futuros

7.1 Conclusões sobre o trabalho

Sabendo da importância das interfaces de interação para a popularização da robótica, e da escassez de estudos que priorizam a intuitividade das mesmas, este trabalho propôs uma metodologia reprodutível a ser empregada na obtenção de vocabulários de gestos intuitivos para interação homem-robô. A metodologia proposta baseia-se na metodologia de [Nielsen et al. \(2004\)](#) e na psicolinguística ([MCNEILL, 1992](#)), no entanto possui algumas características originais, principalmente na fase de experimentos, com destaque para a abordagem baseada na frustração do usuário.

Os estudos realizados a respeito de trabalhos na área de psicolinguística que tratam sobre gestos permitiram concluir que os gestos espontâneos são mais intuitivos. Dessa forma, a metodologia proposta neste trabalho para obtenção de vocabulários de gestos intuitivos toma como base a realização de experimentos baseados em uma abordagem centrada no usuário de forma subconsciente, o que favorece a realização de gestos intuitivos por parte dos voluntários. Além disso, a metodologia propõe o uso de quatro taxas de ocorrência (geral, primeira ocorrência, por voluntário e por tempo) para composição de um Peso Global. É possível que cada uma dessas taxas de ocorrência e o Peso Global funcionem como medidas do grau de intuitividade de um gesto. Esta hipótese porém só poderá ser confirmada quando for realizado um processo de validação do vocabulário obtido.

Ao serem aplicadas, tanto a metodologia proposta neste trabalho quanto a metodologia de Nielsen obtiveram resultados semelhantes, o que indica que a metodologia proposta pode ser considerada consistente, uma vez que ela foi capaz de obter um vocabulário de gestos semelhante ao obtido por uma metodologia já consolidada, a metodologia de Nielsen. Outra razão que faz a metodologia proposta ser interessante é a possibilidade de conceber vocabulários mais complexos, onde mais de um gesto pode ser associado a uma função, o que pode ajudar no aumento da robustez do vocabulário.

Quanto ao aspecto do reconhecimento dos gestos, haja vista que não foram encontrados trabalhos que reconhecem gestos realizados com as duas mãos de maneira simultânea, mesmo não sendo muito elevada, a taxa média de 77,5% conseguida nos testes de reconhecimento pode ser considerada boa. Assim, esse resultado pode ser considerado como um

indicativo para que futuros pesquisadores possam trabalhar mais no desenvolvimento e melhoramento dos modelos de reconhecimento de gestos bimanuais e/ou de corpo inteiro.

Dessa maneira, uma vez que através do uso dos conhecimentos da psicolinguística e da área de HCI, uma metodologia para obter vocabulários de gestos intuitivos foi proposta, os gestos obtidos são consideravelmente distintos (taxa de acertos do reconhecedor de 77,5%) e os resultados se mostraram ser bem promissores, conclui-se que a hipótese apresentada no início deste trabalho foi confirmada. No entanto, visando aprimorar as contribuições dadas, se faz necessário estender os estudos aqui realizados. Dessa maneira, a próxima seção trará a descrição de alguns trabalhos que complementam este, e que devem ser realizados futuramente.

7.2 Trabalhos Futuros

Este trabalho, apesar de trazer contribuições expressivas, deixa algumas lacunas que devem ser fechadas para que se possa ter mais clareza sobre os resultados que a metodologia proposta pode gerar, e consequentemente se tenha um maior desenvolvimento dos temas abordados. Sendo assim, na sequência estão elencados alguns trabalhos que devem ser realizados futuramente objetivando este fim.

7.2.1 Ambiente de realização dos experimentos

Apesar dos experimentos terem sido realizados de maneira satisfatória, percebe-se a necessidade de se ter um ambiente melhor preparado, que permita uma interação ainda mais real entre o voluntário e o robô. Dessa forma, os resultados podem ser mais promissores, e o vocabulário obtido com a aplicação da metodologia proposta pode ser implantado em um ambiente real com uma maior chance de que a interação será realizada com uma taxa de sucesso mais alta.

Um exemplo da necessidade de se ter um ambiente mais adequado para a realização dos experimentos consiste no fato de que por diversas vezes, usuários que estavam realizando experimentos para as funções “Ok” e “Abortar” realizavam gestos mais compatíveis com “parar”, que não estava entre as 7 funções escolhidas, pois como seguiam realizando diversos gestos e o robô continuava a se aproximar deles, começavam a ter receio de serem atropelados pelo robô, mesmo que este estivesse se deslocando a baixas velocidades, e

o pesquisador tendo informado previamente que dispunha de um botão para parada de emergência do robô.

Dessa forma, em um primeiro instante, os experimentos podem ser realizados em um ambiente onde se possa determinar com precisão algumas características desejadas, como por exemplo, a iluminação, o espaço de navegação do robô e o tráfego de pessoas e/ou animais. Posteriormente, o ambiente deve ser o mais próximo do real possível, onde provavelmente ocorrerão alguns eventos não controlados que podem enriquecer ainda mais os resultados dos experimentos.

Além disso, talvez a mera inclusão da função "parar", e incluir na metodologia uma etapa de se mostrar ao voluntário qual o leque de funções abrangido pelo vocabulário, apontando neste leque para qual função ele fará o experimento, já possa ajudar a melhorar os resultados obtidos com os experimentos.

Além dos aspectos levantados aqui, talvez possam ser realizadas outras modificações na metodologia proposta que contribuam para minimizar a chance do voluntário perder o foco durante a realização dos experimentos.

7.2.2 Quantidade e perfil dos voluntários

Por questões de tempo, de locomoção dos materiais utilizados e das autorizações dos direitos de imagem, os experimentos aqui abordados tiveram de ser realizados em universidades, cuja a maioria do voluntários foram da área de engenharias e computação e maiores de 18 anos. No entanto, para que se possa ter uma maior garantia de que o vocabulário obtido com a aplicação da metodologia será intuitivo para o maior número possível de usuários, propõe-se que os experimentos sejam realizados com o máximo possível de voluntários, sendo que os perfis destes sejam tão variados quanto possível.

Assim, uma vez que o objetivo principal deste estudo é contribuir com a popularização da robótica, os experimentos devem ser realizados preferencialmente, no entanto não exclusivamente, com pessoas comuns, que possuam um conhecimento reduzido sobre a robótica e possam usá-la como uma ferramenta para facilitar sua vida, e não como um objeto de estudos, como ocorre na comunidade científica.

Como as crianças e os jovens atuais têm uma tendência maior de se engajarem na tecnologia ([MORAIS, 2002](#)), futuramente serão eles quem possivelmente estarão interagindo com os robôs. Assim, para que se possa desde já iniciar a inserção da robótica como uma

ferramenta cotidiana, se faz necessário um estudo sobre os gestos realizados por crianças e adolescentes, objetivando alcançar uma maior facilidade de interação entre os robôs e tais indivíduos.

7.2.3 Validação do vocabulário

Acredita-se que o vocabulário de gestos obtido com a aplicação da metodologia proposta é intuitivo e robusto. No entanto, tais características só podem ser confirmadas mediante a realização de um processo de validação do vocabulário de gestos obtido.

7.2.3.1 Avaliação do grau de intuitividade do vocabulário de gestos

Ainda em decorrência da falta de tempo destinada a presente pesquisa, não foi possível medir o grau de intuitividade do vocabulário de gestos, tendo isto ficado como extensão deste trabalho. Assim, para realizar tal medição, propõe-se utilizar as medidas de semântica e memorização propostas por [Nielsen et al. \(2004\)](#). No entanto, pelo fato dos gestos aqui obtidos serem, em sua maioria, dinâmicos, a validação deverá ocorrer por meio de vídeos e não de imagens estáticas. Sendo assim, os passos a serem seguidos para poder calcular o Grau de Intuitividade do Vocabulário (GIV) deverão ser os seguintes:

1. **Semântica:** Na abordagem proposta por [Nielsen et al. \(2004\)](#) os usuários recebiam uma lista de gestos e funções, e eles deveriam associar aqueles que eles pensassem ser correspondentes. Desta maneira, era possível calcular o valor semântico do vocabulário por meio da taxa de acertos das correspondências. No entanto, para os gestos encontrados com a metodologia proposta, o ideal é:
 - a) Gravar vídeos curtos, contendo gestos realizados por um ator. Essa gravação deve ser feita de maneira a garantir o anonimato do ator, para que os aspectos faciais não influenciem na correspondência de cada gesto.
 - b) Os vídeos devem ser passados para várias pessoas (aconselha-se que esse procedimento seja realizado na forma de formulários *on-line*) e elas devem responder para cada gesto, qual função ele melhor representa. Nesta etapa é importante inserir gestos que não estejam presentes no vocabulário e também uma resposta do tipo “Não corresponde a nenhuma função”, a fim de oferecer a cada partici-

pante a possibilidade real de expressar o que ele pensa de cada gesto, e que o resultado da aplicação da validação não acabe sendo direcionado.

- c) Após o fim da coleta das respostas, deve-se calcular a o Grau de Semântica do Vocabulário (GSV), o qual pode ser feito por meio da seguinte expressão:

$$GSV = \frac{NA}{NA + NE} \quad (15)$$

onde:

NA: Representa número de acertos. Aqui devem ser considerados todos os gestos, até aqueles que não correspondem a nenhuma função do vocabulário.

NE: Representa o número de erros. Neste caso, deve-se contabilizar quantas vezes um gesto é dito representar uma função que ele não representa, quando um gesto válido é dito não representar nenhuma função, ou quando um falso gesto é dito representar uma das funções.

2. **Memorização:** O desenvolvimento desta etapa é bem semelhante àquela proposta por [Nielsen et al. \(2004\)](#), no entanto, ao invés de utilizar imagens, serão utilizados vídeos. Assim, de posse dos vídeos utilizados para o cálculo de GSV, devem-se realizar testes de memorização com voluntários, os quais devem ser executados da seguinte maneira:

- a) Descrever para o voluntário todas as possíveis funções que o robô deve realizar.
- b) Mostrar para o voluntário vídeos que representem os gestos que correspondem a cada uma das funções apresentadas. Como os gestos são supostamente intuitivos, o mesmo vídeo deve ser passado no máximo duas vezes para cada voluntário.
- c) Após isso, deve-se mostrar para o voluntário o nome de uma função, e este deve realizar instantaneamente o gesto que ele pensa ser correspondente a tal função. Para isso, o voluntário deve ser instruído previamente a realizar o gesto que ele observou anteriormente em um dos vídeos, e que ele pensa corresponder à função apresentada. Assim, ao ser mostrada uma função, deve-se medir o tempo até o início da realização dos gestos e gravar quais foram os gestos realizados.
- d) De posse das gravações dos gestos realizados e dos seus respectivos tempos de realização, deve-se calcular o Grau de Memorização do Vocabulário (GMV) por meio da seguinte expressão:

$$GMV = \frac{\sum_{i=1}^{NV} \sum_{j=1}^{NGV_i} g(i, j)}{\sum_{i=1}^{NV} NGV_i}. \quad (16)$$

onde:

$$g(i, j) = h(i, j)f(T_{ij}), \quad (17)$$

$$h(i, j) = \begin{cases} 1, & \text{se o } i\text{'ésimo voluntário acertou o } j\text{'ésimo gesto} \\ 0, & \text{se o } i\text{'ésimo voluntário errou o } j\text{'ésimo gesto} \end{cases}, \quad (18)$$

$$f(t) = \begin{cases} 1, & \text{se } t < t_{min} \\ 1 - \left(\frac{t - t_{min}}{t_{max} - t_{min}} \right), & \text{se } t_{min} \leq t \leq t_{max} \\ 0, & \text{se } t > t_{max} \end{cases} \quad (19)$$

Sendo:

t_{min} : tempo considerado ideal para início do gesto (a ser determinado).

t_{max} : tempo máximo admissível para início do gesto (a ser determinado).

NV : o número de voluntários que realizaram os testes de memorização.

NGV_i : O número de gestos realizados pelo i 'ésimo voluntário.

T_{ij} : O tempo de que o i 'ésimo voluntário levou para iniciar o j 'ésimo gesto.

Sendo assim, o GIV pode ser calculado pela a média aritimética entre os dois valores calculados anteriormente:

$$GIV = \frac{GSV + GMV}{2} \quad (20)$$

7.2.3.2 Experimentos práticos com o vocabulário obtido

Para poder chegar a conclusões ainda mais completas sobre o vocabulário de gestos obtido com a aplicação da metodologia proposta, além de medir seu grau de intuitividade, é necessária uma etapa adicional de experimentos envolvendo um robô e situações concretas. Assim, será possível observar a efetividade de cada um dos gestos presentes no vocabulário e levantar dados que, após analisados, possam melhorar a metodologia proposita. Para isso, essa validação pode ser feita da seguinte maneira:

- Criar um ambiente semelhante ao real e dotar o robô da capacidade de realizar tarefas reais.

- Dotar o robô da capacidade de reconhecer os gestos presentes no vocabulário.
- Instruir voluntários a interagir com o robô por meio de gestos. Essas instruções devem ser dadas de tal maneira que os voluntários possam executar qualquer gesto que eles pensem ser ideais para a comunicação, sem influenciá-los a realizar apenas os gestos presentes no vocabulário.
- Todo o processo de interação deve ser gravado e analisado.
- Após todo o processo de análise, deve-se levantar estatísticas sobre a interação, como por exemplo: taxa de acertos, percentual de gestos diferentes daqueles presentes no vocabulário e quantidade de gestos diferentes realizados antes de acertar o gesto que o robô realmente conhece.
- Assim, fazendo uso das estatísticas calculadas, deve-se levantar conclusões e questionamentos que possam, respectivamente, validar o vocabulário obtido e melhorar a efetividade da metodologia proposta.

7.2.4 Compartilhamento dos resultados

Para melhorar o desenvolvimento dos estudos sobre gestos intuitivos para HRI, é necessário que os gestos e os resultados das avaliações sejam compartilhados de maneira mais efetiva. Assim, propõe-se a criação de uma plataforma *on-line* onde seja possível fazer este compartilhamento dos dados obtidos pelas aplicações da metodologia proposta, bem como os dados de trabalhos que utilizam outras metodologias, mas que se destinam ao mesmo objetivo. Dessa maneira, existirá um local de consulta para quem deseja obter vocabulários de gestos intuitivos para a interação homem-robô ou para quem deseja dados para desenvolver estudos sobre reconhecedores de gestos. Posteriormente os dados obtidos com os experimentos, bem como a metodologia proposta e o vocabulário de gestos obtido por ela serão disponibilizados *online*.

7.2.5 Integração entre fala e gestos

Mesmo os gestos constituindo uma ótima interface para a HRI, para que a robótica seja efetivamente popularizada, a interface de interação deve ser ainda mais completa e intuitiva. Logo, como foi abordado várias vezes por McNeill (1992), fala e gestos estão diretamente ligados. Assim, necessita-se de um estudo bem elaborado sobre a integração

efetiva entre fala e gestos, objetivando a possibilidade da complementação da interface de interação entre homem e robô.

7.2.6 Inferência de contextos em ambientes interacionais

Como foi abordado por [Wachs \(2006\)](#), existe a necessidade dos vocabulários de gestos serem o mais complexos possíveis. Essa complexidade está ligada à relação entre gestos e funções. Sendo assim, um gesto pode representar mais de uma função e uma função pode ser representada por mais de um gesto. Logo, mesmo que o presente trabalho tenha proposto uma metodologia que possibilita encontrar vocabulários de gestos complexos, ele não resolve o problema de, dada uma determinada função, determinar em qual momento um gesto é mais intuitivo que outro, ou então qual a função que um determinado gesto está representando em um determinado momento.

Sendo assim, a solução para esses problemas consiste em determinar qual o atual contexto em que o gesto está inserido no momento da sua execução. Logo, propõe-se também, como trabalho futuro, a realização de um estudo detalhado sobre a inferência efetiva dos possíveis contextos presentes em um determinado ambiente interacional.

7.2.7 Melhoramento do modelo de reconhecimento dos gestos

Mesmo considerando 77,5% uma taxa de reconhecimento boa, ela pode ser aumentada por meio de alguns melhoramentos no modelo de reconhecimento. Para isso, deve-se desenvolver algum mecanismo que possibilite melhorar a qualidade dos gestos tomados como modelos de comparação. Outra ação importante a ser realizada é fazer os testes de reconhecimento com um quantitativo maior de dados gestuais, uma vez que a quantidade de dados utilizada neste trabalho não foi a ideal. Além disso, pode-se ainda, estudar a possibilidade de adaptar outro modelo de reconhecimento, para tentar melhorar a taxa de reconhecimento de gestos realizados com as duas mãos e/ou de corpo inteiro.

7.2.8 Correlação entre trajetória, velocidade, número de repetições e características únicas de um gesto

Após a análise dos gestos, percebeu-se a presença de três fatores, além da trajetória do movimento, que influenciam diretamente na diferenciação de alguns gestos: a velocidade do movimento, o número de repetições e as suas características únicas.

Como o modelo de reconhecimento utilizado não levou essas informações em consideração, alguns gestos, como o “mais rápido”, por exemplo, foram confundidos. Muitas vezes, quando as trajetórias são semelhantes, o que realmente pode diferenciar alguns gestos são essas três informações. Dessa maneira, um modelo de reconhecimento, para ser mais eficaz, deve fornecer a possibilidade de acrescentar a velocidade do movimento, o número de repetições e algumas características específicas de cada gesto como informações adicionais.

No entanto, ainda não se sabe como essas informações se correlacionam, por isso fica difícil criar um reconhecedor baseado nesse conjunto de informações. Por isso, um trabalho muito importante a ser realizado futuramente, é o estudo e determinação da correlação entre essas quatro variáveis que podem representar um gesto: trajetória, velocidade, número de repetições, e características únicas (posição dos dedos e da cabeça, contorções no corpo, etc.). Dessa maneira, seria possível criar um modelo de representação de gestos mais eficiente para o processo de reconhecimento.

7.2.9 Interface de interação baseada em gestos universais

Muitos trabalhos de HRI como, ([KIM; KWAK; CH, 2006](#); [ENDE et al., 2011](#); [BODIROŽA; STERN; EDAN, 2012](#)), apontam que os gestos podem oferecer uma interface de comunicação entre homem e robô eficiente e intuitiva, no entanto existe ainda a necessidade de se tentar encontrar vocabulários de gestos universais para que a interação entre homem e robô possa ser realizada por qualquer pessoa, sem que exista o problema da influência da cultura, apontada por [Archer \(1997\)](#) e [McNeill \(1992\)](#), sobre a interface de interação.

Em seu trabalho, [Stern et al. \(2008\)](#) diz não acreditar na existência de gestos universais, no entanto, [Kendon \(2007\)](#) crê que exista um padrão para os gestos e que as pessoas, mesmo de diferentes culturas, compartilham esse conhecimento. Assim, no trabalho apresentado por [Goldin-Meadow et al. \(2008\)](#), concluiu-se, através de experiências

envolvendo pessoas de diferentes países, culturas e línguas, que na ausência de fala a ordem dos gestos é a mesma, o que trás um indício da possibilidade de existência de vocabulários de gestos capazes de compor uma interface de comunicação, que além de intuitiva, possa ser utilizada na interação com um robô em qualquer parte do mundo.

Neste sentido, seria de fundamental importância a realização de um trabalho envolvendo pesquisadores de diversos países, nos quais os idiomas nativos tenham uma estrutura significativamente diferente, permitindo assim, que se possa investigar se diferenças culturais e linguísticas afetam significativamente os vocabulários de gestos obtidos para um mesmo conjunto de funções.

Referências

- AKHTARUZZAMAN Md; SHAFIE A. A. Evolution of Humanoid Robot and contribution of various countries in advancing the research and development of the platform. In: *Control Automation and Systems (ICCAS), 2010 International Conference*. [S.l.]: IEEE, 2010. p. 1021 – 1028. Citado na página [20](#).
- ALON, J. et al. Simultaneous localization and recognition of dynamic hand gestures. *Proceedings - IEEE Workshop on Motion and Video Computing, MOTION 2005*, p. 254–260, 2007. Citado na página [84](#).
- ALY, A.; TAPUS, A. Speech to Head Gesture Mapping in Multimodal Human-Robot Interaction. *Service Orientation in Holonic and Multi-Agent Manufacturing Control*, p. 183–196, 2012. Citado na página [24](#).
- ANGLE, C. *Genghis, a Six legged Autonomous Walking Robot*. Tese (Doutorado) — Massachusetts Institute of Technology, 1989. Citado na página [21](#).
- ARCHER, D. Unspoken Diversity : Cultural Differences in Gestures. *Qualitative Sociology*, v. 20, n. 1, p. 79–105, 1997. ISSN 01620436. Citado na página [93](#).
- ASIMOV, I. *I, Robot*. [S.l.]: Gnome Press, 1950. 253 p. Citado na página [24](#).
- AWARD, G.; SUTHERLAND, a. Automatic Skin Segmentation for Gesture Recognition Combining Region and Support Vector Machine Active Learning. *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, Ieee, p. 237–242, 2006. Citado na página [84](#).
- BARATTINI, P.; MORAND, C.; ROBERTSON, N. A proposed gesture set for the control of industrial collaborative robots. In: IEEE (Ed.). *RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. [S.l.]: IEEE, 2012. p. 132–137. ISBN 978-1-4673-4606-1. Citado 6 vezes nas páginas [11](#), [39](#), [40](#), [41](#), [62](#) e [104](#).
- BETTIO, R. W. D. et al. Model and implementation of body movement recognition using Support Vector Machines and Finite State Machines with Cartesian coordinates input for gesture-based interaction. *JCS&T*, v. 13, n. 2, p. 69–75, 2013. Citado na página [83](#).
- BHOKSE, B.; KARWANKAR, A. R. Hand Gesture recognition Using Neural Network. v. 2, n. 1, p. 1–8, 2015. Citado na página [84](#).
- BHUYAN, M. K. FSM-based recognition of dynamic hand gestures via gesture summarization using key video object planes. p. 724–735, 2012. Citado na página [83](#).
- BISWAS, K. K.; BASU, S. K. Gesture recognition using Microsoft Kinect®. *The 5th International Conference on Automation, Robotics and Applications*, Ieee, v. 2, p. 100–103, dec 2011. Citado na página [84](#).
- BLUETHMANN, W. et al. Robonaut : A Robot Designed to Work with Humans in Space . *Autonomous Robot 14*, p. 179–197, 2003. Citado na página [22](#).
- BOBICK, A. F.; SOCIETY, I. C. A State-Based Approach to the Representation and Recognition of Gesture. v. 19, n. 12, p. 1325–1337, 1997. Citado na página [83](#).

- BODIROŽA, S.; STERN, H. I.; EDAN, Y. Dynamic gesture vocabulary design for intuitive human-robot dialog. In: *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction - HRI '12*. [S.l.: s.n.], 2012. p. 111. ISBN 9781450310635. ISSN 21672121. Citado na página 93.
- BOWDEN, R. et al. A Linguistic Feature Vector for the Visual Interpretation of Sign Language. In: *Computer Vision-ECCV*. [S.l.]: Springer Berlin Heidelberg, 2004. p. 390–401. Citado na página 84.
- BREAZER, C.; ARYANANDA, L. Recognition of Affective Communicative Intent in Robot-Directed Speech. v. 12, n. 1, p. 83–104, 2002. Citado na página 24.
- BREMNER, P. et al. Conversational Gestures in Human-Robot Interaction. In: *Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics*. [S.l.]: IEEE, 2009. p. 1645–1649. ISBN 9781424427949. Citado na página 26.
- BROCCIA, G.; LIVESU, M.; SCATENI, R. Gestural Interaction for Robot Motion Control. In: ABATE, A. F.; NAPPI, M.; TORTORA, G. (Ed.). *Eurographics Italian Chapter Conference (2011)*. [S.l.]: University of Cagliari, 2011. p. 61–66. Citado na página 39.
- BROOKS, R. A. Intelligence Without Reason. n. 1293, 1991. Citado na página 21.
- BURKE, J. et al. Final Report for the DARPA/NSF Interdisciplinary Study on Human–Robot Interaction. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, v. 34, n. 2, p. 103–112, may 2004. Citado 3 vezes nas páginas 15, 43 e 104.
- BUSS, M. et al. Towards Proactive Human-Robot Interaction in Human Environments. In: *Cognitive Infocommunications (CogInfoCom)*. [S.l.: s.n.], 2011. p. 1–6. Citado na página 24.
- CAPEK, K. et al. Rossum’s universal robots. *Prague, CZ*, 1920. Citado na página 20.
- CHEN, Y.-T.; TSENG, K.-T. Multiple-angle Hand Gesture Recognition by Fusing SVM Classifiers. *2007 IEEE International Conference on Automation Science and Engineering*, Ieee, p. 527–530, sep 2007. Citado na página 84.
- CHEVALLEREAU, C. et al. *Bipedal robots: modeling, design and walking synthesis*. 78. ed. [S.l.]: John Wiley & Sons, 2010. Citado na página 20.
- CHOI, E. et al. Towards successful user interaction with systems: Focusing on user-derived gestures for smart home systems. *Applied Ergonomics*, Elsevier Ltd, v. 45, n. 4, p. 1196–1207, 2014. ISSN 00036870. Citado 2 vezes nas páginas 53 e 54.
- CORADESCHI, S. et al. GiraffPlus: Combining social interaction and long term monitoring for promoting independent living. *2013 6th International Conference on Human System Interactions (HSI)*, Ieee, p. 578–585, jun 2013. Citado na página 22.
- CORP, M. *MSN - Microsoft Developer Network*. 2016. Disponível em: <<https://msdn.microsoft.com/en-us/library/hh973073.aspx>>. Acesso em: 07/02/2016. Citado 2 vezes nas páginas 11 e 81.

- COSTALONGA, T. O. et al. Gesture-Based Controllers to Guide a Quadrotor Using Kinect Sensor. *2014 Joint Conference on Robotics: SBR-LARS Robotics Symposium and Robocontrol*, p. 109–112, 2014. Citado na página [41](#).
- COVER, T.; HART, P. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, v. 13, n. 1, p. 21–27, 1967. ISSN 0018-9448. Citado na página [79](#).
- DAHL, T.; BOULOS, M. Robots in Health and Social Care: A Complementary Technology to Home Care and Telehealthcare? *Robotics*, v. 3, n. 1, p. 1–21, dec 2013. ISSN 2218-6581. Citado na página [22](#).
- DAHLBÄCK, N.; JÖNSSON, A.; AHRENBERG, L. Wizard of Oz studies - why and how. *Knowledge-Based Systems*, v. 6, n. 4, p. 258–266, 1993. ISSN 09507051. Citado 2 vezes nas páginas [48](#) e [53](#).
- DAVIDSON, W. L. Definition of Intuition. *Mind*, v. 7, n. 26, p. 304–310, 1882. Citado na página [38](#).
- DAVIS, J.; SHAH, M. *Visual gesture recognition*. 1994. Citado na página [84](#).
- DRESZER, D. et al. Autonomization of Two-Wheeled Inverted Pendulum Robot – Kinect-Based Obstacle Detection and Avoidance System. 2011. Citado na página [64](#).
- DROESCHEL, D. et al. Towards Robust Mobility , Flexible Object Manipulation , and Intuitive Multimodal Interaction for Domestic Service Robots. In: *RoboCup 2011: Robot Soccer World Cup XV*. [S.l.: s.n.], 2012. p. 51–62. Citado na página [43](#).
- DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern classification*. [S.l.]: John Wiley & Sons, 2012. Citado na página [82](#).
- ELMEZAIN, M.; AL-HAMADI, A. A Hidden Markov Model-Based Isolated and Meaningful Hand Gesture Recognition. p. 156–163, 2009. Citado na página [84](#).
- ELMEZAIN, M. et al. A Hidden Markov Model-Based Continuous Gesture Recognition System for Hand Motion Trajectory. p. 3–6, 2008. Citado na página [84](#).
- ENDE, T. et al. A human-centered approach to robot gesture based communication within collaborative working processes. *IEEE International Conference on Intelligent Robots and Systems*, p. 3367–3374, 2011. ISSN 2153-0858. Citado na página [93](#).
- ENGELBERGER, J. F. Robotics in the 21st Century. *Scientific American*, v. 273, n. 3, p. 166, 1995. Citado na página [20](#).
- FUJII, T.; LEE, J. H.; OKAMOTO, S. Gesture Recognition System for Human-Robot Interaction and Its Application to Robotic Service Task. In: *Proceedings of the International MultiConference of Engineers and Computer Scientists 2014 Vol I, IMECS 2014*. Hong Kong: [s.n.], 2014. I, p. 1–6. Citado 5 vezes nas páginas [11](#), [39](#), [44](#), [62](#) e [104](#).
- GALLOWAY, J. C.; RYU, J.-C.; AGRAWAL, S. K. Babies driving robots: self-generated mobility in very young infants. *Intelligent Service Robotics*, v. 1, n. 2, p. 123–134, jan 2008. ISSN 1861-2776. Citado na página [24](#).

- GAUS, Y. F. A.; WONG, F. Hidden Markov Model-Based Gesture Recognition with Overlapping Hand-Head/Hand-Hand Estimated Using Kalman Filter. *2012 Third International Conference on Intelligent Systems Modelling and Simulation*, Ieee, p. 262–267, feb 2012. Citado na página 84.
- GOLDIN-MEADOW, S. et al. The natural order of events: how speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences of the United States of America*, v. 105, n. 27, p. 9163–8, jul 2008. ISSN 1091-6490. Citado 2 vezes nas páginas 25 e 93.
- GOSLING, J. et al. *The Java Language Specification Java SE 8 Edition*. [S.l.], 2015. Citado na página 30.
- HASANUZZAMAN, M. et al. Face and Gesture Recognition Using Subspace Method for Human-Robot Interaction. In: *Advances in Multimedia Information Processing-PCM 2004*. [S.l.]: Springer Berlin Heidelberg, 2005. p. 369–376. Citado 3 vezes nas páginas 26, 62 e 104.
- HASSANI, A. Z. *Touch versus in-air Hand Gestures : Evaluating the acceptance by seniors of Human-Robot Interaction using Microsoft Kinect*. Tese (Doutorado) — Dissertação (Master in Computer Science) - University of Twente, 2011. Citado na página 39.
- HONG, P. et al. Gesture Modeling and Recognition Using Finite State Machines. In: *IEEE Conference on Face and Gesture Recognition*. [S.l.]: IEEE, 2000. p. 1–6. Citado na página 84.
- HSE. *Safety signs and signals : The Health and Safety (Safety Signs and Signals) Guidance on Regulations*. 2. ed. [S.l.]: HSE Books, 2009. 44 p. Citado na página 41.
- JAKOB, N. Why you only need to test with 5 users. *Jakob Nielsen's Alterbox*. [URL: www.useit.com/alterbox/20000319.html], v. 19, 2000. Citado na página 62.
- JONES, J. L. Robots at The Tipping Point: The road to the iRobot Roomba. *Robotics & Automation Magazine*, n. March, p. 76–78, 2006. Citado na página 22.
- KATALINIC, E. B. EVALUATION OF KINECT DEPTH SENSOR FOR USE IN MOBILE ROBOTICS. v. 22, n. 1, p. 147–148, 2011. Citado na página 64.
- KENDON, A. *Gesture: Visible action as utterance*. [S.l.]: Cambridge University Press, 2004. 412 p. Citado 2 vezes nas páginas 39 e 40.
- KENDON, A. An Agenda for Gesture Studies. *Semiotic review of books*, v. 7, n. 3, p. 1–22, 2007. Citado 2 vezes nas páginas 25 e 93.
- KIM, J. An HMM-based threshold model approach for gesture recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 21, n. 10, p. 961–973, 1999. ISSN 01628828. Citado na página 84.
- KIM, J. W.; NAM, T. J.; PARK, T. CompositeGesture: Creating Custom Gesture Interfaces with Multiple Mobile or Wearable Devices. *International Journal on Interactive Design and Manufacturing*, p. 1–6, 2014. ISSN 19552513. Citado na página 45.

KIM, K. K. K. K. K.; KWAK, K. C. K. K. C.; CH, S. Y. C. S. Y. Gesture analysis for human-robot interaction. *2006 8th International Conference Advanced Communication Technology*, v. 3, p. 1824–1827, 2006. Citado na página [93](#).

KIM, Y. C.; JANG, J. S. Remote Gesture Control Using HMM Classifier Based on Electric Potential Sensors. v. 10, n. 2, p. 413–420, 2015. Citado na página [84](#).

KLASSNER, F.; ANDERSON, S. D. LEGO MindStorms : Not Just for K-12 Anymore. v. 2001, 2001. Citado na página [21](#).

KORTENKAMP, D.; HUBER, E.; BONASSO, R. P. Recognizing and interpreting gestures on a mobile robot The proximity space method Recognizing gestures. In: *Proceedings of the National Conference on Artificial Intelligence*. [S.l.: s.n.], 1996. p. 915–921. Citado 2 vezes nas páginas [62](#) e [104](#).

KOTA, S. R. et al. Principal Component Analysis for Gesture Recognition Using SystemC. *2009 International Conference on Advances in Recent Technologies in Communication and Computing*, Ieee, p. 732–737, 2009. Citado na página [84](#).

KRISTOFFERSSON, A.; CORADESCHI, S.; LOUTFI, A. A Review of Mobile Robotic Telepresence. *Advances in Human-Computer Interaction*, v. 2013, p. 1–17, 2013. ISSN 1687-5893. Citado na página [22](#).

KULK, J.; WELSH, J. A Low Power Walk for the NAO Robot. 2005. Citado na página [22](#).

LEE, C. Online, interactive learning of gestures for human/robot interfaces. In: *Proceedings of IEEE International Conference on Robotics and Automation*. [S.l.]: IEEE, 1996. v. 4, p. 2982–2987. ISBN 0-7803-2988-0. Citado na página [24](#).

LEE, I.; CHA, J.; KWON, O. Multi-modal Interaction System for Smart TV Environments. *2014 IEEE International Symposium on Multimedia*, Ieee, p. 263–264, dec 2014. Citado na página [45](#).

LEE, J. et al. Hand region extraction and gesture recognition from video stream with complex background through entropy analysis. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, v. 2, p. 1513–6, jan 2004. ISSN 1557-170X. Citado na página [84](#).

LIU, K. et al. Multi-HMM classification for hand gesture recognition using two differing modality sensors. *2014 IEEE Dallas Circuits and Systems Conference (DCAS)*, Ieee, p. 1–4, oct 2014. Citado na página [84](#).

MASHOOD, A.; NOURA, H. A Gesture Based Kinect for Quadrotor Control. p. 298–301, 2015. Citado na página [39](#).

MASTER, R. M.; STOCKHOLM, S. T. Robot Obstacle Avoidance using the Kinect. 2011. Citado na página [64](#).

MATUSZEK, C. et al. Learning from Unscripted Deictic Gesture and Language for Human-Robot Interactions. *Proceedings of AAAI 2014*, 2014. Citado 2 vezes nas páginas [11](#) e [36](#).

MCNEILL, D. *Hand and mind: What gestures reveal about thought*. Chicago/Illinois: The University of Chicago Press, 1992. 416 p. Citado 12 vezes nas páginas 11, 25, 35, 38, 39, 46, 52, 74, 76, 85, 91 e 93.

MEDEIROS, A. C.; TAVARES, T. A.; FONSECA, I. E. da. How to design an user interface based on gestures? In: *Design, User Experience, and Usability: Design Discourse*. [S.l.]: Springer, 2015. p. 63–74. Citado 3 vezes nas páginas 11, 50 e 51.

MEENA, R.; JOKINEN, K.; WILCOCK, G. Integration of gestures and speech in human-robot interaction. In: *IEEE 3rd International Conference on Cognitive Infocommunications (CogInfoCom)*. [S.l.]: IEEE, 2012. p. 673–678. ISBN 978-1-4673-5188-1. Citado 4 vezes nas páginas 11, 39, 40 e 62.

MICHAELIS. *Dicionário Online - Dicionários Michaelis - UOL*. 2009. Disponível em: <<http://michaelis.uol.com.br/>>. Acesso em: 06/02/2016. Citado na página 37.

MICROSOFT. *Kinect for Windows Sensor Components and Specifications*. 2015. Disponível em: <<https://msdn.microsoft.com/en-us/library/jj131033.aspx>>. Acesso em: 22/12/2015. Citado 3 vezes nas páginas 11, 30 e 63.

MITRA, S.; ACHARYA, T. Gesture Recognition : A Survey. *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS*, v. 37, n. 3, p. 311–324, 2007. Citado na página 76.

MORAIS, J. F. R. de. *Filosofia da Ciência e da Tecnologia*. 1. ed. [S.l.]: Papirus, 2002. 148 p. ISBN 85-308-0490-2. Citado na página 87.

MORAN, M. E. Evolution of robotic arms. *Journal of robotic surgery*, v. 1, n. 2, p. 103–111, jan 2007. ISSN 1863-2483. Citado na página 20.

MUKAI, T. et al. Development of a nursing-care assistant robot RIBA that can lift a human in its arms. *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Ieee, p. 5996–6001, oct 2010. Citado na página 22.

NETO, A. N. R. et al. Development of interactive virtual environments using java and kinect technologies. In: ACM. *Proceedings of the 19th Brazilian symposium on Multimedia and the web*. [S.l.], 2013. p. 11–12. Citado na página 64.

NICOLELIS, M. A. L.; CHAPIN, J. K. Controlling Robots with the Mind. *SCIENTIFIC AMERICAN-AMERICAN EDITION*, v. 287, n. 4, p. 46–55, 2002. Citado na página 24.

NIELSEN, M. et al. A procedure for developing intuitive and ergonomic gesture interfaces for HCI. *Gesture-Based Communication in Human-Computer Interaction*, Springer Berlin Heidelberg, p. 409–420, 2004. ISSN 03029743. Citado 11 vezes nas páginas 26, 46, 47, 49, 50, 51, 52, 53, 85, 88 e 89.

OBAID, M. et al. A Framework for User-Defined Body Gestures to Control a Humanoid Robot. *International Journal of Social Robotics*, v. 6, n. 3, p. 383–396, 2014. ISSN 18754805. Citado 2 vezes nas páginas 39 e 104.

OHTA, T. Intuitive gesture set on touch screen for ad-hoc file sharing among multiple mobile devices. *2014 IEEE 3rd Global Conference on Consumer Electronics (GCCE)*, Ieee, p. 751–752, oct 2014. Citado na página 45.

ONISHI, M. et al. Generation of Human Care Behaviors by Human-Interactive Robot RI-MAN. *Proceedings 2007 IEEE International Conference on Robotics and Automation*, Ieee, p. 3128–3129, apr 2007. ISSN 1050-4729. Citado na página 22.

ORACLE, A. I. *Oracle*. 2015. Disponível em: <<http://www.oracle.com/index.html>>. Acesso em: 22/12/2015. Citado na página 30.

ORACLE, C. *MySQL*. 2015. Disponível em: <<https://www.mysql.com/products-enterprise/techspec.html>>. Acesso em: 22/12/2015. Citado na página 30.

ORACLE, C. *NetBeans IDE 8.0.2*. 2015. Disponível em: <<https://netbeans.org/community/releases/80/relnotes.html>>. Acesso em: 22/12/2015. Citado na página 30.

PAUL, P.; PINKY, A. *ADHD Powered - Learning with Laughter*. 2013. Disponível em: <<https://adhdpowered.wordpress.com/tag/wilson-memorial-hospital/>>. Citado 2 vezes nas páginas 11 e 37.

PEREIRA, F. G. *Detecção de Pessoas e Reconhecimento de Gestos para Interação e Cooperação entre Seres Humanos e Robôs*. Tese (Doutorado) — Thasis (PhD in engineer electrical - Federal University of Espírito Santo - UFES, 2012. Citado 5 vezes nas páginas 11, 20, 42, 43 e 63.

POWERS, A. What robotics can learn from HCI. *Interactions*, v. 15, n. 2, p. 67, 2008. ISSN 10725520. Citado 2 vezes nas páginas 26 e 46.

PRESS, W. *Iconic Gestures boost Language Acquisition*. 2012. Disponível em: <<https://susipucci.wordpress.com/2012/01/22/iconic-gestures-boost-language-acquisition/>>. Citado 2 vezes nas páginas 11 e 35.

RAIBERT, M. et al. BigDog , the Rough-Terrain Quadruped Robot. p. 6–9, 2008. Citado na página 21.

RASHID, O.; AL-HAMADI, A.; MICHAELIS, B. A framework for the integration of gesture and posture recognition using HMM and SVM. *2009 IEEE International Conference on Intelligent Computing and Intelligent Systems*, Ieee, v. 1, p. 572–577, nov 2009. Citado na página 84.

ROBOT, A. M. *Pioneer 3-DX*. [S.l.], 2011. 2 p. Acesso em: 22/12/2015. Citado na página 30.

ROBOTCUP. *Robot Cup Home*. 2011. Disponível em: <<http://www.robocup2011.org/en/>>. Acesso em: 24/12/2015. Citado na página 43.

ROGALLA, O. et al. Using gesture and speech control for commanding a robot assistant. *Proceedings. 11th IEEE International Workshop on Robot and Human Interactive Communication*, Ieee, p. 454–459, 2002. Citado na página 24.

SAKAGAMI, Y.; WATANABE, R.; AOYAMA, C. The intelligent ASIMO: System overview and integration. n. October, 2002. Citado na página 21.

SANTOS, C. C.; NUNES, M. A. S. N. Prospecção tecnológica de reconhecimento de gestos através de dados obtidos com o kinect. v. 11, p. 1–15, 2015. Citado 2 vezes nas páginas 63 e 64.

SIEGWART, R.; NOURBAKHSH, I. R.; SCARAMUZZA, D. *Introduction to Autonomous Mobile Robots second edition*. [S.l.]: MIT Press, 2011. ISBN 9780262015356. Citado na página 21.

SIEKMANN, J.; HARTMANIS, J.; LEEUWEN, J. V. Gesture-Based Communication in Human-Computer Interaction. In: *Lecture Notes in Artificial Intelligence*. [S.l.]: Springer, 1999. ISBN 3540669353. Citado na página 39.

STARNER, T.; PENTLAND, A. S. Real-Time American Sign Language Recognition Hidden Markov Models from Video Using. *AAAI Technical Report FS-96-05*, p. 109–116, 1996. Citado na página 84.

STEFAN, A. et al. Translation and scale-invariant gesture recognition in complex scenes. *Proceedings of the 1st ACM international conference on Pervasive Technologies Related to Assistive Environments - PETRA '08*, ACM Press, New York, New York, USA, p. 1, 2008. Disponível em: <<http://portal.acm.org/citation.cfm?doid=1389586.1389595>>. Citado na página 84.

STERN, H. I.; WACHS, J. P.; EDAN, Y. Optimal hand gesture vocabulary design using psycho-physiological and technical factors. *FGR 2006: Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition*, v. 2006, p. 257–262, 2006. Citado na página 26.

STERN, H. I. et al. Optimal consensus intuitive hand gesture vocabulary design. *Proceedings - IEEE International Conference on Semantic Computing 2008, ICSC 2008*, Ieee, p. 96–103, aug 2008. Citado 2 vezes nas páginas 26 e 93.

TEE, K. P. et al. Gesture-based attention direction for a telepresence robot: Design and experimental study. In: *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. [S.l.: s.n.], 2014. p. 4090–4095. ISBN 978-1-4799-6934-0. ISSN 21530866. Citado na página 104.

TMSUK. *Tmsuk Robots*. 2009. Robots p. Disponível em: <<http://www.tmsuk.co.jp/english/robots.html>>. Acesso em: 17/02/2015. Citado na página 22.

TOKYO, C.; NARA, Y.-k. Training AIBO like a Dog. In: *13th International Workshop on Robot and Human Interactive Communication*. [S.l.: s.n.], 2004. Citado na página 21.

TRAN, T.-H. et al. Dynamic hand gesture recognition using RGB-D motion history and kernel descriptor. *2014 International Conference on Advanced Technologies for Communications (ATC 2014)*, Ieee, p. 268–273, oct 2014. Citado na página 84.

URIBE, A.; ALVES, S. Task planning for Human-Robot Interaction. *Robotics Symposium*, ..., 2011. Citado na página 26.

VERMA, R.; DEV, A. Vision based hand gesture recognition using finite state machines and fuzzy logic. *2009 International Conference on Ultra Modern Telecommunications & Workshops*, IEEE, p. 1–6, oct 2009. Citado na página 84.

VOLPE, R. et al. Rocky 7: A next generation mars rover prototype. *Advanced Robotics*, v. 11, n. 4, p. 341–358, 1996. Citado na página 21.

WACHS, J. *Optimal Hand Gesture Vocabulary Design Methodology for Virtual Robotic Control*. 289 p. Tese (Doutorado) — Ben-Gurion University of the Negev, 2006. Citado 8 vezes nas páginas [11](#), [46](#), [47](#), [49](#), [50](#), [53](#), [72](#) e [92](#).

WALDHERR, S. et al. Template-Based Recognition of Pose and Motion Gestures On a Mobile Robot. In: *AAAI/IAAI*. [S.l.: s.n.], 1998. p. 977–982. Citado 5 vezes nas páginas [11](#), [26](#), [39](#), [62](#) e [104](#).

WANG, X. et al. Hidden-Markov-Models-Based Dynamic Hand Gesture Recognition. *Mathematical Problems in Engineering*, v. 2012, p. 1–11, 2012. ISSN 1024-123X. Citado na página [84](#).

WEXELBLAT, A. Research Challenges in Gesture: Open Issues and Unsolved Problems. In: WACHSMUTH, I.; FRÖHLICH, M. (Ed.). *Gesture and Sign Language in Human-Computer Interaction*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998, (Lecture Notes in Computer Science, v. 1371). cap. 1, p. 1–11. ISBN 978-3-540-64424-8. Citado 3 vezes nas páginas [25](#), [38](#) e [45](#).

WOBBROCK, J. O.; MORRIS, M. R.; WILSON, A. D. User-defined gestures for surface computing. *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*, p. 1083, 2009. ISSN 1605582468. Citado na página [49](#).

WU, H.; WANG, J.; ZHANG, X. Combining hidden Markov model and fuzzy neural network for continuous recognition of complex dynamic gestures. *The Visual Computer*, Springer Berlin Heidelberg, 2015. Citado na página [84](#).

WU, H.; WANG, J.; ZHANG, X. L. User-centered gesture development in TV viewing environment. *Multimedia Tools and Applications*, v. 75, p. 733–760, 2016. Citado na página [49](#).

YAMATO, J.; OHYA, J.; ISHII, K. *Recognizing human action in time-sequential images using hidden Markov model*. [S.l.]: IEEE, 1992. 379–385 p. Citado na página [84](#).

YANG, H.-d.; PARK, A.-y.; LEE, S.-w. Gesture Spotting and Recognition for Human – Robot Interaction. In: *IEEE Transactions on Robotics*. [S.l.]: IEEE, 2007. v. 23, n. 2, p. 256–270. Citado 7 vezes nas páginas [11](#), [26](#), [44](#), [45](#), [62](#), [83](#) e [104](#).

YEASIN, M.; CHAUDHURI, S. Visual understanding of dynamic hand gestures. *Pattern Recognition*, v. 33, n. 11, p. 1805–1817, nov 2000. ISSN 00313203. Citado na página [84](#).

YOHANAN, S.; MACLEAN, K. E. The Role of Affective Touch in Human-Robot Interaction: Human Intent and Expectations in Touching the Haptic Creature. *International Journal of Social Robotics*, v. 4, n. 2, p. 163–180, dec 2011. ISSN 1875-4791. Citado 4 vezes nas páginas [11](#), [26](#), [41](#) e [42](#).

ZHANG, T. et al. Semantic Gesture Recognition Based on Cognitive Behavioral Model. *2014 International Conference on Information Science & Applications (ICISA)*, Ieee, n. 1, p. 1–4, may 2014. Citado na página [84](#).

Apêndice A – Descrição detalhada da aplicação da metodologia

Esse apêndice objetiva trazer com uma maior riqueza de detalhes como a metodologia proposta foi aplicada.

A.1 Etapa 1: Escolha das funções

Como o trabalho não objetivou a obtenção de um vocabulário para um ambiente de interação específico, as funções escolhidas foram baseadas em trabalhos sobre HRI que apresentam os gestos utilizados e/ou as funções realizadas por um robô móvel. Assim, na tabela 4 podem ser vistos os gestos escolhidos e os trabalhos utilizados na análise.

Tabela 4 – Funções escolhidas e os trabalhos onde elas aparecem.

Função	Trabalhos
ABORTAR	(BARATTINI; MORAND; ROBERTSON, 2012)
APONTAR	(KORTENKAMP; HUBER; BONASSO, 1996) (HASANUZZAMAN et al., 2005) (BURKE et al., 2004) (TEE et al., 2014)
ATENÇÃO	(BARATTINI; MORAND; ROBERTSON, 2012) (YANG; PARK; LEE, 2007)
MAIS DEVAGAR	(BARATTINI; MORAND; ROBERTSON, 2012) (OBAID et al., 2014)
MAIS RÁPIDO	(BARATTINI; MORAND; ROBERTSON, 2012) (OBAID et al., 2014)
OK	(BARATTINI; MORAND; ROBERTSON, 2012) (BURKE et al., 2004)
SIGA-ME	(WALDHERR et al., 1998) (BARATTINI; MORAND; ROBERTSON, 2012) (FUJII; LEE; OKAMOTO, 2014)

A.2 Etapa 2: Captura dos dados Gestuais

A.2.1 Logística utilizada no ambiente

Para realizar a captura dos dados gestuais, o ambiente foi escolhido de tal forma que um voluntário não pudesse ver o experimento de outro, pois, como a metodologia baseia-se na frustração, ao perceber que um voluntário se frustrou, outro voluntário poderia tentar realizar gestos que não sejam aqueles que ele pensava ser intuitivos para uma determinada função. Assim, como normalmente os experimentos foram realizados em salas e ou laboratórios, foi possível garantir o sigilo destes, fazendo com que outros voluntários esperassem sua vez em uma área externa ao ambiente de realização dos mesmos. Um exemplo do ambiente utilizado nos experimentos pode ser visto na Figura 23.

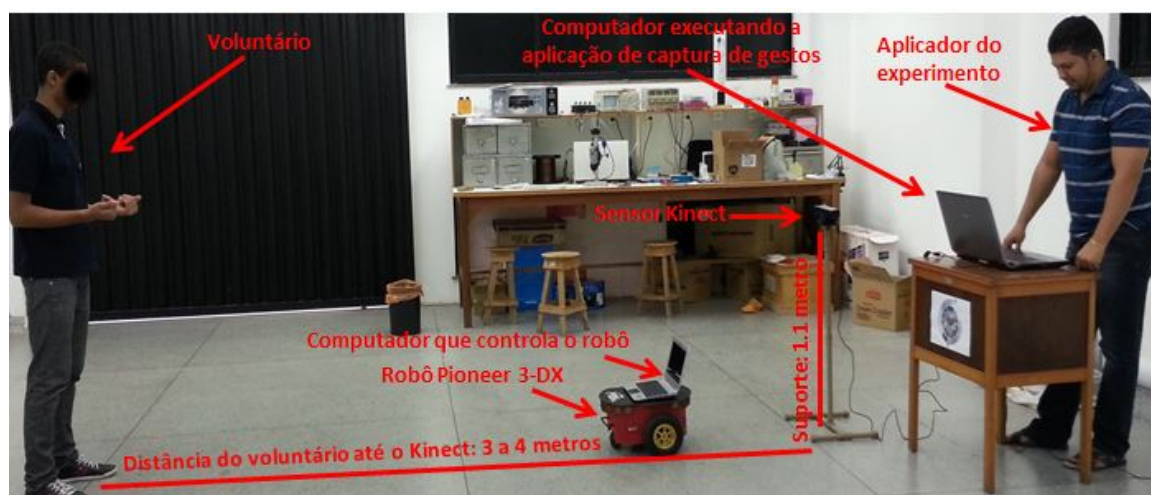


Figura 23 – Ambiente utilizado

A.2.2 Etapas de Realização dos Experimentos

No momento da realização dos experimentos cada voluntário, independente da função a qual lhe era designada, passava pelo mesmo processo, o qual pode ser dividido em algumas etapas.

A.2.2.1 Apresentação

- Olá, nós somos do grupo de Pesquisa em Robótica da UFS e estamos realizando alguns experimentos para um trabalho de mestrado. O trabalho consiste em estudar interfaces intuitivas baseadas em gestos para a interação homem robô.
- Qual o seu nome?

- Então <**nome do voluntário**>, antes de explicar o funcionamento do experimento, vamos lhe mostrar os nossos equipamentos.
- Eis o nosso robô móvel com o qual você irá interagir.
- Temos também o sensor *kinect*, o qual irá capturar os seus dados gestuais para que possam ser analisados posteriormente.
- Dessa maneira, antes de qualquer coisa, gostaria que você lê-se e assinasse esse termo de consentimento e esse formulário, os quais, além de trazer informações sobre os experimentos, também serve para que você autorize, ou não, o aparecimento de suas imagens em futuras publicações e para oferecer alguns dados estatísticos para o nosso estudo (entrega o termo de consentimento e o formulário para o voluntário, e espera ele ler e assinar).
- Para iniciar a explicação, gostaria que você se posicionasse aqui (indica o ponto onde o voluntário deveria permanecer, 3 a 4 metros do *Kinect*). Esse será o local onde você deverá ficar para poder interagir com o robô.

A.2.2.2 Orientações

Com o voluntário posicionado no local designado para a realização dos experimentos, o mesmo era instruído a interagir com o robô. Para cada uma das sete funções escolhidas tinha-se uma instrução diferente, a saber:

1. Roteiro para a função “Mais rápido”

- O robô se deslocará na sua direção com uma certa velocidade. Assim, você deve pedir para ele vir mais rápido, ou seja, que ele aumente a sua velocidade. No entanto, o robô reconhece apenas um gesto para a função “mais rápido” e você deverá adivinhar qual é esse gesto. Caso você realize um gesto e o robô continue se deslocando com a mesma velocidade, significa que o gesto que você realizou não é aquele que ele reconhece como o correspondente da função “mais rápido”, então você pode tentar outro gesto até que ele reconheça algum e aumente sua velocidade.
- Você perceberá que o robô reconheceu o seu gesto quando ele, bruscamente, aumentar a velocidade de deslocamento. No entanto, não se preocupe, pois, caso ele não reconheça o seu gesto, ele irá parar antes de tocar em você.

2. Roteiro para a função “Mais Devagar”

- O robô virá na sua direção com uma certa velocidade. Assim, você deve pedir para ele vir mais devagar, ou seja, que ele diminua a velocidade. No entanto, o robô reconhece apenas um gesto para a função “mais devagar” e por isso você deve adivinhar qual é esse gesto. Caso você realize o gesto e o robô continue se deslocando com a mesma velocidade, significa que o gesto realizado não é

aquele que o robô reconhece então, você pode tentar outro gesto até que ele reconheça algum e diminua sua velocidade.

- Você perceberá que o robô reconheceu o seu gesto quando ele, bruscamente, diminuir a sua velocidade de deslocamento. No entanto, não se preocupe, pois, caso ele não reconheça o seu gesto, ele irá parar antes de tocar em você.

3.Roteiro para a função “Apontar”

- O robô estará parado e você deverá apontar para algum lugar no ambiente para o qual você deseja que ele se desloque. No entanto, o robô reconhece apenas um gesto para a função apontar e por isso você deve adivinhar qual é esse gesto. Caso o robô continue parado, significa que o gesto que você realizou não corresponde ao gesto que ele reconhece como o representante da função “apontar”. Dessa maneira, caso você realize um gesto e o robô não reconheça, você pode tentar outros gestos até que ele reconheça algum e se desloque para o ponto apontado.

4.Roteiro para a função “Abortar”

- Imagine que o robô está realizando uma tarefa, a qual é vir na sua direção. Você quer que ele aborte essa tarefa, ou seja, que ele deixe a execução da tarefa no ponto em que ela está. Então, enquanto ele se desloca na sua direção, você deve realizar gestos que sinalizem o comando de “abortar”. No entanto, o robô só reconhece um gesto para esta função, por isso você deve adivinhar qual é esse gesto. Caso o robô continue vindo em sua direção, significa que ele não reconheceu seu comando e por isso não abortou a tarefa. Sendo assim, você deverá realizar gestos que você pensa representar essa função até que o robô reconheça algum e aborte a execução da tarefa.
- Você perceberá que o robô reconheceu o seu gesto quando ele parar. No entanto, não se preocupe, pois caso ele não reconheça nenhum dos seus gestos, ele irá parar antes de tocar em você.

5.Roteiro para a função “Ok”

- Imagine que o robô está realizando uma tarefa, a qual é vir na sua direção. Você quer passar para ele que a tarefa foi realizada com sucesso. Então, enquanto ele se desloca na sua direção, você deve realizar gestos que sinalizem o comando de “ok”. No entanto, o robô só reconhece um gesto para esta função, e por isso você deve adivinhar qual é esse gesto. Caso o robô continue vindo em sua direção, significa que ele não reconheceu seu comando. Sendo assim, você deverá realizar gestos que você pensa representar essa função até o que robô reconheça algum.
- Você perceberá que o robô reconheceu o seu gesto quando ele parar. No entanto, não se preocupe, pois caso ele não reconheça nenhum dos seus gestos, ele irá parar antes de tocar em você.

6. Roteiro para a função “Atenção”

- O robô estará parado e você deverá chamar a atenção dele realizando gestos. No entanto, o robô só reconhece um gesto para a função “atenção” e por isso você deve adivinhar qual é esse gesto. Caso o robô continue parado, significa que o gesto que você realizou não corresponde ao gesto que ele reconhece como o representante da função “atenção”. Dessa maneira, você pode tentar outros gestos até que ele reconheça algum.
- Você perceberá que o robô reconheceu o seu gesto quando ele se mover para frente.

7. Roteiro para a função “Siga-me”

- O robô estará parado e você deverá pedir para ele te seguir. Assim, enquanto ele estiver parado você deve realizar gestos que peçam para ele te seguir. No entanto, o robô só reconhece um gesto para a função “siga-me” e por isso você deve adivinhar qual é esse gesto. Caso o robô continue parado, significa que o gesto que você realizou não corresponde ao gesto que ele reconhece como o representante da função “siga-me”. Dessa maneira, você pode tentar outros gestos até que ele reconheça algum.
- Você perceberá que o robô reconheceu o seu gesto quando ele se mover na sua direção.

A.2.2.3 Calibragem do *Kinect*

O sensor *kinect* necessita reconhecer um voluntário na cena para poder capturar os dados do seu esqueleto virtual. Para isso o voluntário era instruído da seguinte maneira:

- Para que o *kinect* possa usado, ele deve ser calibrado. Para isso, quando eu sinalizar, você, vagarosamente, virá na direção do *kinect* com os braços levantados e realizando movimentos de um lado para o outro. Entendido?
- Ok. Pode vir.

Neste momento a rotina de reconhecimento do esqueleto virtual era acionada na aplicação de captura e esperava-se até que a aplicação retornasse uma indicação de que o esqueleto virtual foi reconhecido. Caso não fosse possível reconhecer o esqueleto, pedia-se para o voluntário voltar para a posição inicial e realizar o movimento novamente. Nos experimentos realizados, não houveram casos em que o voluntário teve que repetir o movimento mais de duas vezes.

- Pronto, o *Kinect* já foi calibrado. Agora você pode voltar para a sua posição inicial para que o experimento possa iniciar.
- Alguma dúvida? (esclarecer possíveis dúvidas do voluntário)

A.2.2.4 Demonstração

Após explicar para o voluntário o que ele deveria realizar, tentava-se fazer com que ele acreditasse que o robô realmente interagiria com ele por meio de gestos, para isso seguia-se o seguinte roteiro:

- Como eu lhe disse, o robô reconhece vários gestos, no entanto apenas um gesto corresponde à função *<nome da função designada ao voluntário>*. Dessa maneira, como nós estamos conversando e realizando gestos, o robô teve o seu reconhecedor de gestos desativado, pois, assim não corremos o risco de realizar um gesto e ele sair realizando funções no ambiente. Logo, para testarmos o reconhecedor, existe um gesto predefinido que está associado à função “recuar”, que é o seguinte: (realiza o gesto ilustrado na figura 24).
- Por favor, realize o gesto para eu ver se você entendeu. (espera o voluntário realizar o gesto).
- Muito bem. Sendo assim, para você ver que o robô está realmente funcionando, quando eu sinalizar, você deverá realizar o gesto de recuar. Dessa maneira, o robô irá recuar 50 centímetros, até se aproximar do suporte onde o *Kinect* se encontra. Então, quando o robô terminar o recuo, você já pode realizar o gesto que corresponde à função *<nome da função designada ao voluntário>*.
- Entendido? (caso o voluntário não entenda, explica-se novamente).



Figura 24 – Gesto de recuar

A.2.2.5 Início do experimento

Para iniciar o experimento, o pesquisador posicionava-se atrás do computador, que estava executando a aplicação de captura de gestos e onde estava localizado o controle remoto, e dizia o seguinte para o voluntário:

- O reconhecedor de gesto do robô foi habilitado, agora ele já reconhece seus gestos.
- Então, quando eu der o sinal, você irá executar o gesto de recuo que eu te expliquei antes. Quando o robô recuar os 50 centímetros e parar você já poderá realizar os gestos que você pensa corresponderem à função de *<função designada ao voluntário>*.
- Entendido? (esclarecer possíveis dúvidas levantadas pelo voluntário)
- Então, pode pedir para o robô recuar.

Aqui o pesquisador ficava olhando para a tela e ao ver que o voluntário tinha realizado o gesto de recuar, ele enviava, via controle remoto, um comando para o robô recuar 50 centímetros. Após o robô recuar os 50 centímetros, caso a função exigisse que o robô se movimentasse, o pesquisador acionava o comando correspondente à velocidade a ser utilizada na função.

Assim, após iniciar a tarefa, ao se deslocar entre 2,5 e 3 metros o robô parava e os dados capturados eram armazenados em disco. Caso o pesquisador percebesse que a frustração tinha sido alcançada antes do tempo estipulado, ele acionava um comando para o robô parar e armazenava os dados capturados pelo *Kinect*.

A utilização do movimento de recuo serviu para que o voluntário pensasse que o robô realmente interagira por meio de gestos. Dessa maneira, foi possível fazer com que ele acreditasse que no momento da realização do experimento, o robô não estava respondendo aos gestos realizados porque nenhum desses gestos correspondia à função designada ao robô. No entanto, como o robô não reconhecia nenhum tipo de gesto e não respondia a nenhum comando que não fosse dado pelo controle remoto que estava com o pesquisador, ele realiza a ação de recuar, não porque tinha reconhecido o gesto de recuo realizado pelo voluntário, e sim porque o pesquisador pressionara um botão no seu controle remoto.

Isso também ocorreu no momento da interação, pois, uma vez que o robô só respondia aos comandos vindos do controle remoto, ele nunca respondia adequadamente aos comandos dados por meio dos gestos realizados pelo voluntário. Sendo assim, como o voluntário acreditava que o robô reconhecia gestos e que o robô realmente interagira, como o robô não respondia aos gestos realizados, o voluntário inconscientemente sentia a necessidade de realizar outros gestos. Assim, no decorrer do experimento, todas as ideias de gestos intuitivos para aquela função eram tentadas sem sucesso e o voluntário se frustrava por não poder interagir. Tal frustração podia ser percebida pelo pesquisador, por meio de expressões faciais ou corporais realizadas pelo voluntário.

Após o término do experimento, como o *kinect* capturou todos os gestos realizados e os mesmos foram armazenados, foi possível realizar, posteriormente, a etapa de análise dos gestos.

A.2.3 Observação sobre as etapas realizadas

No fim de cada experimento, o voluntário sempre perguntava qual era o gesto que representava aquela função. Então, era respondido que, por questões de sigilo científico, não era possível divulgar essa informação, mas que os resultados e explicações viriam a ser conhecidos por meio de futuras publicações.

Essa informação não era dada pois um voluntário poderia simplesmente comentar com outro que ainda não havia realizado o experimento, e dessa maneira esse outro voluntário provavelmente não realizaria o experimento da maneira adequada. Por esse motivo, no fim de cada experimento solicitava-se ao voluntário que ele não divulgasse como o experimento funcionava, a fim de fazer com que todos os voluntários tivessem o mesmo conhecimento sobre o experimento no momento da aplicação. Tabela de Gestos Realizados (TGR)

A.2.4 Etapa 3: Análise dos dados gestuais

Seguindo a etapa 3 da metodologia, para analisar os vídeos armazenados anteriormente foi criada uma aplicação JAVA que reproduz esses arquivos de vídeo. Essa aplicação permite pausar a reprodução e também avançar e retroceder quadros. Dessa maneira, todo o processo de análise foi realizado por meio da observação de cada um dos gestos realizados em todos os vídeos. É importante destacar, que uma única pessoa foi responsável por todo esse processo de observações e que nada foi feito de maneira automática. Desta forma, cada um dos gestos foi analisado de maneira muito cautelosa para que os conhecimentos do observador não interferissem nos resultados. Foram necessárias quase 18 horas de análise, o que resultou nas duas tabelas listadas abaixo.

- Tabela de Descrição dos Gestos (TDG): (Tabela 5 do apêndice B)
- Tabela de Gestos Realizados (TGR): (Tabela 6 do apêndice C)

A.2.5 Etapa 4: Escolha do vocabulário

A partir das tabelas criadas na etapa anterior, seguindo os 5 primeiros passos da metodologia proposta, foi possível chegar às cinco tabelas listadas abaixo.

- Tabela com a Taxa de Ocorrência Geral (TOG) dos gestos capturados e seus respectivos Pesos: (Tabela 7 do apêndice D)
- Tabela com a Taxa de Primeira Ocorrência (TPO) dos gestos capturados e seus respectivos Pesos: (Tabela 8 do apêndice D)
- Tabela com a Taxa de Ocorrência por Voluntário (TOV) dos gestos capturados e seus respectivos Pesos: (Tabela 9 do apêndice D)
- Tabela com a Taxa de Ocorrência por Tempo (TOT) dos gestos capturados e seus respectivos Pesos: (Tabela 10 do apêndice D)
- Tabela com Peso Geral (PG) de cada um dos gestos capturados: (Tabela 11 do apêndice E)

Desta maneira, de acordo com o sexto passo da metodologia, analisando a tabela com Peso Geral, foi possível chegar ao vocabulário representado na Figura 25.

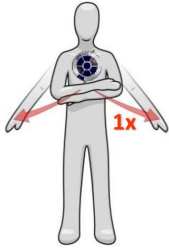
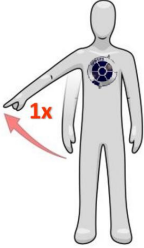
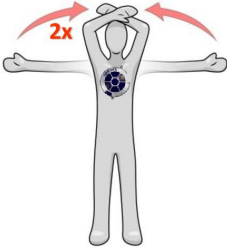



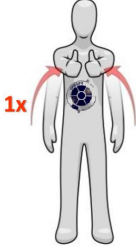


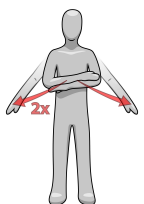
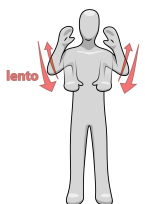
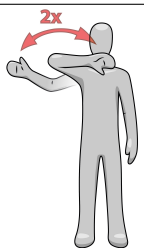
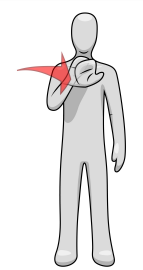
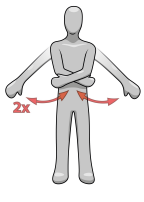
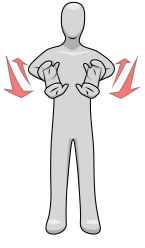
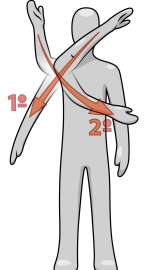
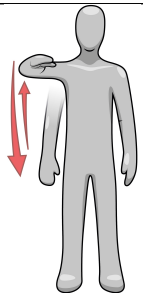
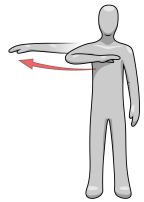
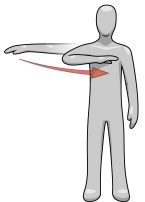
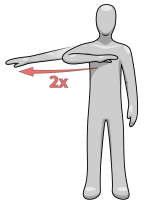
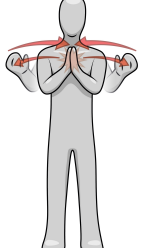
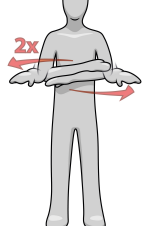
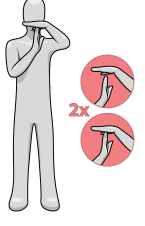
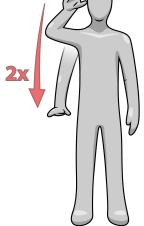
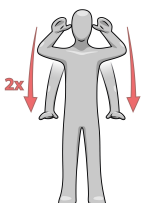
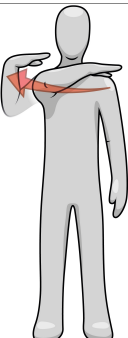
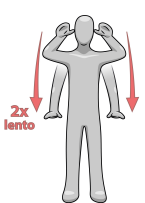
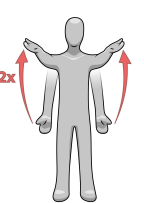
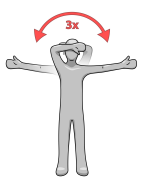
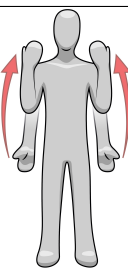
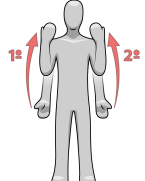
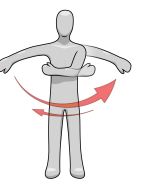
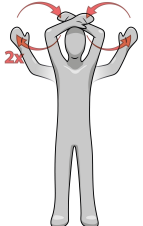

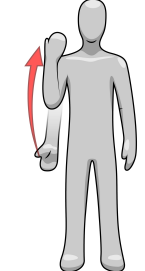

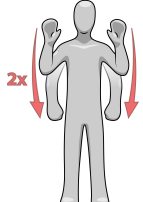
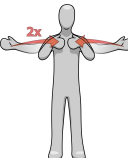
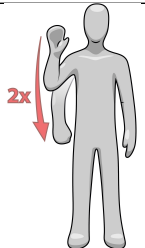
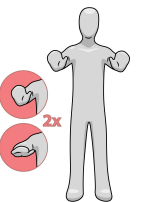
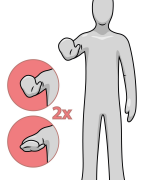
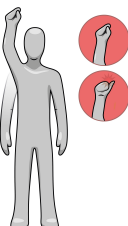

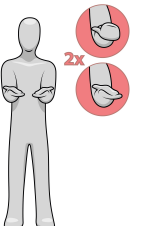
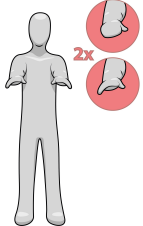
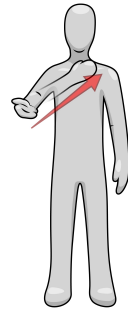
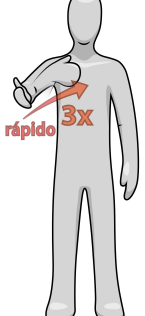
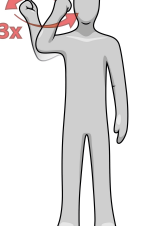
Abortar	Apontar	Atenção	Mais Devagar
			
Mais Rápido	Ok		Siga-me
			

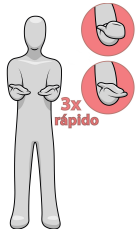
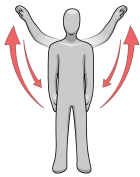
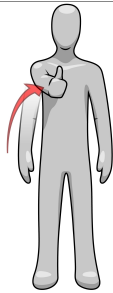
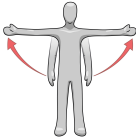
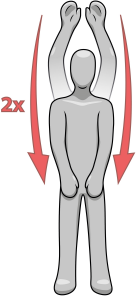
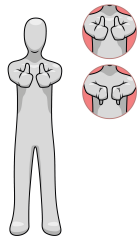
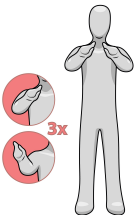

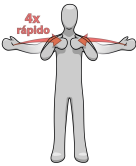

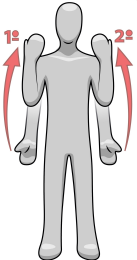
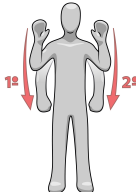
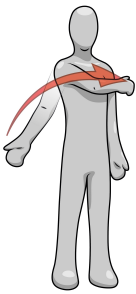
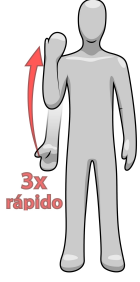
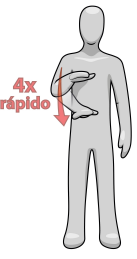
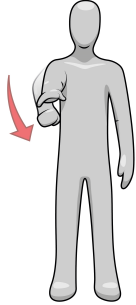
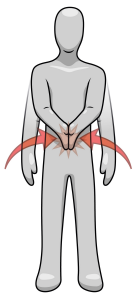
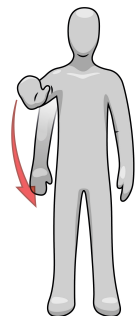
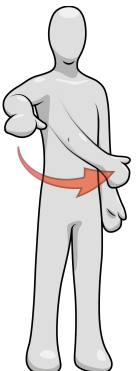
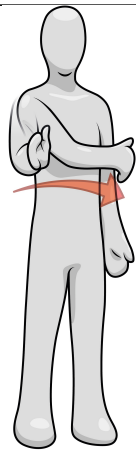
Figura 25 – Vocabulário obtido com a aplicação da metodologia proposta

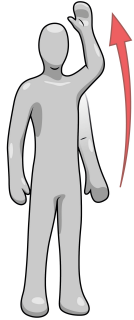
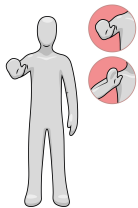
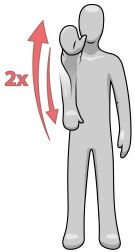

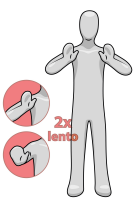
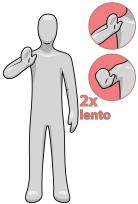
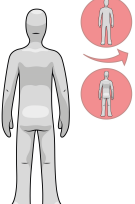
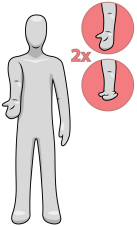
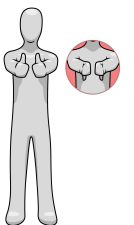
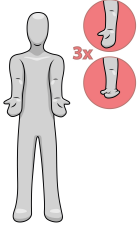
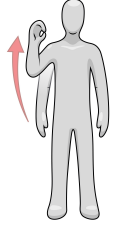
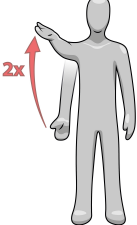
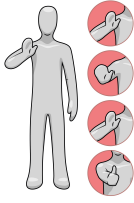

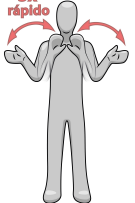
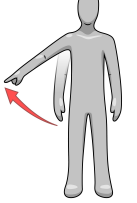
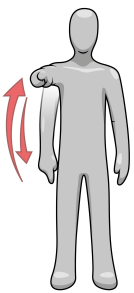
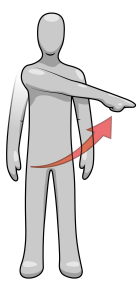
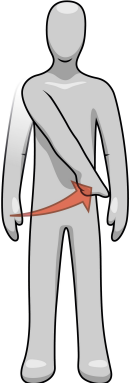
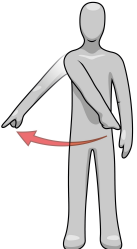
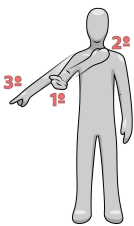
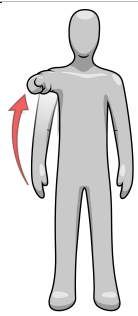
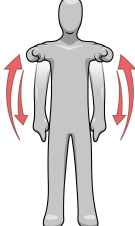
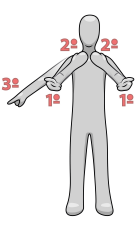
Apêndice B – Descrição dos gestos obtidos nos experimentos

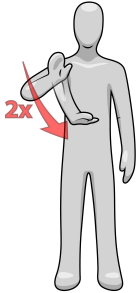
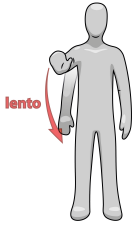
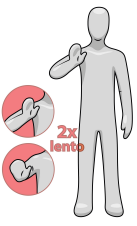
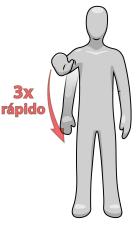

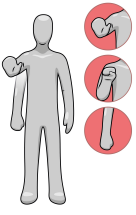
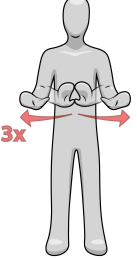
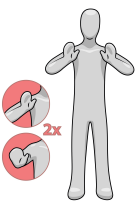
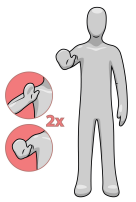
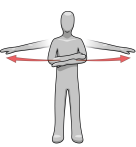

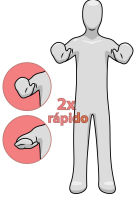
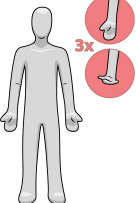
Tabela 5 – Tabela de Descrição dos Gestos (TDG): Tabela onde todos os gestos encontrados com a aplicação dos experimentos são identificados e descritos

ID	Gesto	ID	Gesto	ID	Gesto	ID	Gesto
1		2		3		4	
5		6		7		8	
9		10		11		12	
13		14		15		16	

ID	Gesto	ID	Gesto	ID	Gesto	ID	Gesto
17		18		19		20	
21		22		23		24	
25		26		27		28	
29		30		31		32	
33		34		35		36	
37		38		39		40	

ID	Gesto	ID	Gesto	ID	Gesto	ID	Gesto
41		42		43		44	
45		46		47		48	
49		50		51		52	
53		54		55		56	
57		58		59		60	

ID	Gesto	ID	Gesto	ID	Gesto	ID	Gesto
61		62		63		64	
65		66		67		68	
69		70		71		72	
73		74		75		76	
77		78		79		80	
81		82		83		84	

ID	Gesto	ID	Gesto	ID	Gesto	ID	Gesto
85		86		87		88	
89		90		91		92	
93		94		95		96	
97							

Apêndice C – Gestos de cada voluntário em cada tarefa.

Tabela 6 – Tabela de Gestos Realizados (TGR): Todos os gestos realizados por cada um dos voluntários para cada tarefa.

Função	Voluntário	Gesto	Tempo
Abortar	1	1	1
Abortar	1	1	2
Abortar	1	5	3
Abortar	1	2	4
Abortar	1	3	5
Abortar	1	2	7
Abortar	1	6	8
Abortar	1	6	10
Abortar	1	4	11
Abortar	1	5	13
Abortar	1	3	14
Abortar	1	2	15
Abortar	1	4	16
Abortar	1	2	18
Abortar	1	4	19
Abortar	1	7	20
Abortar	2	8	1
Abortar	2	2	3
Abortar	2	5	5
Abortar	2	9	7
Abortar	2	10	9
Abortar	2	5	11
Abortar	2	10	13
Abortar	2	11	14
Abortar	2	8	15
Abortar	2	8	17
Abortar	2	12	20
Abortar	3	5	3
Abortar	3	2	6
Abortar	3	2	8
Abortar	3	2	13

Função	Voluntário	Gesto	Tempo
Abortar	3	2	16
Abortar	4	4	1
Abortar	4	13	2
Abortar	4	11	15
Abortar	5	2	1
Abortar	5	2	2
Abortar	5	2	3
Abortar	5	2	7
Abortar	5	2	9
Abortar	5	14	12
Abortar	5	2	13
Abortar	5	2	14
Abortar	5	3	19
Abortar	5	2	21
Abortar	5	3	23
Abortar	6	2	3
Abortar	6	2	4
Abortar	6	2	5
Abortar	6	15	12
Abortar	6	2	15
Abortar	7	2	1
Abortar	7	2	2
Abortar	7	5	4
Abortar	7	16	5
Abortar	7	2	7
Abortar	7	5	9
Abortar	7	5	10
Abortar	7	5	11
Abortar	7	2	12
Abortar	7	17	13
Abortar	7	17	14
Abortar	7	2	15
Abortar	7	5	18
Abortar	8	2	1
Abortar	8	18	3
Abortar	8	2	4
Abortar	8	2	7

Função	Voluntário	Gesto	Tempo
Abortar	8	18	10
Abortar	8	2	12
Abortar	8	2	15
Abortar	8	18	17
Abortar	8	3	19
Abortar	8	3	21
Abortar	9	19	1
Abortar	9	19	6
Abortar	9	19	10
Abortar	9	2	11
Abortar	9	3	15
Abortar	9	17	16
Abortar	9	16	17
Abortar	9	16	19
Abortar	9	19	20
Abortar	9	3	21
Abortar	9	17	22
Abortar	10	5	1
Abortar	10	2	2
Abortar	10	5	8
Abortar	10	5	11
Abortar	10	3	13
Abortar	10	2	20
Abortar	10	5	22
Abortar	10	3	24
Abortar	10	5	27
Abortar	11	2	1
Abortar	11	3	1
Abortar	11	2	4
Abortar	11	3	6
Abortar	11	2	8
Abortar	11	3	11
Abortar	11	2	13
Abortar	11	2	15
Abortar	11	2	16
Abortar	11	5	17
Abortar	11	5	18

Função	Voluntário	Gesto	Tempo
Abortar	11	2	21
Abortar	11	3	22
Abortar	11	3	25
Abortar	12	8	2
Abortar	12	2	3
Abortar	12	18	3
Abortar	12	12	5
Apontar	1	76	1
Apontar	1	76	3
Apontar	1	76	6
Apontar	1	76	7
Apontar	1	76	9
Apontar	1	76	11
Apontar	1	76	13
Apontar	1	76	14
Apontar	1	76	16
Apontar	1	76	19
Apontar	1	77	20
Apontar	1	76	21
Apontar	2	78	1
Apontar	2	78	5
Apontar	2	78	8
Apontar	2	78	12
Apontar	2	78	14
Apontar	2	78	16
Apontar	2	78	19
Apontar	2	78	21
Apontar	3	76	3
Apontar	3	79	7
Apontar	3	76	13
Apontar	3	79	18
Apontar	3	76	22
Apontar	4	80	2
Apontar	4	31	4
Apontar	4	3	7
Apontar	4	5	18
Apontar	4	59	24

Função	Voluntário	Gesto	Tempo
Apontar	5	81	2
Apontar	5	31	8
Apontar	5	27	12
Apontar	5	81	13
Apontar	5	77	20
Apontar	5	27	25
Apontar	5	31	26
Apontar	6	27	3
Apontar	6	27	5
Apontar	6	77	8
Apontar	6	78	13
Apontar	6	78	21
Apontar	6	78	25
Apontar	7	76	2
Apontar	7	78	6
Apontar	7	76	8
Apontar	7	76	13
Apontar	7	76	19
Apontar	7	76	23
Apontar	8	76	1
Apontar	8	81	3
Apontar	8	81	4
Apontar	8	27	5
Apontar	8	5	10
Apontar	8	37	13
Apontar	8	79	14
Apontar	8	27	16
Apontar	8	26	20
Apontar	8	79	21
Apontar	8	27	24
Apontar	8	5	25
Apontar	9	82	1
Apontar	9	76	4
Apontar	9	82	7
Apontar	9	83	9
Apontar	9	82	12
Apontar	9	79	16

Função	Voluntário	Gesto	Tempo
Apontar	9	84	17
Apontar	9	76	19
Apontar	9	76	22
Apontar	9	82	24
Apontar	10	76	3
Apontar	10	76	5
Apontar	10	76	7
Apontar	10	76	9
Apontar	10	76	11
Apontar	10	76	16
Apontar	10	76	20
Apontar	10	76	23
Apontar	10	76	26
Apontar	11	78	3
Apontar	11	77	8
Apontar	11	77	13
Apontar	11	76	15
Apontar	11	76	20
Apontar	12	76	1
Apontar	12	76	4
Apontar	12	78	6
Apontar	12	78	11
Apontar	12	76	14
Apontar	12	76	17
Apontar	12	79	20
Atenção	1	21	1
Atenção	1	22	3
Atenção	1	23	6
Atenção	1	21	11
Atenção	1	24	14
Atenção	1	21	19
Atenção	1	25	22
Atenção	1	26	25
Atenção	1	26	26
Atenção	2	27	1
Atenção	2	21	3
Atenção	2	21	8

Função	Voluntário	Gesto	Tempo
Atenção	2	28	13
Atenção	2	27	18
Atenção	3	28	1
Atenção	3	21	8
Atenção	3	28	11
Atenção	4	21	1
Atenção	4	29	2
Atenção	4	28	4
Atenção	4	30	5
Atenção	4	28	7
Atenção	4	29	9
Atenção	4	31	10
Atenção	4	28	12
Atenção	4	29	15
Atenção	4	29	17
Atenção	4	28	20
Atenção	4	2	24
Atenção	4	31	26
Atenção	5	34	1
Atenção	5	35	2
Atenção	5	22	3
Atenção	5	28	4
Atenção	5	27	10
Atenção	5	26	12
Atenção	5	26	15
Atenção	6	32	2
Atenção	6	32	3
Atenção	6	33	5
Atenção	6	33	7
Atenção	6	36	9
Atenção	6	33	11
Atenção	6	32	13
Atenção	6	33	17
Atenção	6	32	18
Atenção	6	33	21
Atenção	6	33	22
Atenção	6	36	23

Função	Voluntário	Gesto	Tempo
Atenção	6	33	25
Atenção	7	26	1
Atenção	7	4	5
Atenção	7	26	6
Atenção	7	27	8
Atenção	7	4	11
Atenção	7	11	14
Atenção	7	27	15
Atenção	7	1	17
Atenção	7	27	19
Atenção	7	27	21
Atenção	7	4	24
Atenção	8	26	1
Atenção	8	37	10
Atenção	8	26	12
Atenção	8	37	15
Atenção	8	37	17
Atenção	8	26	18
Atenção	8	5	22
Atenção	8	5	26
Atenção	9	37	1
Atenção	9	22	2
Atenção	9	31	3
Atenção	9	29	6
Atenção	9	38	9
Atenção	9	29	13
Atenção	9	38	15
Atenção	9	22	19
Atenção	9	31	23
Atenção	9	28	26
Atenção	9	28	29
Atenção	10	4	3
Atenção	10	28	5
Atenção	10	28	6
Atenção	10	42	10
Atenção	10	28	15
Atenção	10	28	17

Função	Voluntário	Gesto	Tempo
Atenção	10	27	20
Atenção	10	26	24
Atenção	10	26	26
Atenção	10	26	28
Atenção	10	28	30
Atenção	11	28	1
Atenção	11	28	3
Atenção	11	22	7
Atenção	11	4	10
Atenção	11	4	12
Atenção	11	28	15
Atenção	11	28	17
Atenção	11	25	23
Atenção	11	28	25
Atenção	11	28	30
Atenção	12	28	2
Atenção	12	4	5
Atenção	12	4	8
Atenção	12	28	11
Atenção	12	22	14
Atenção	12	4	19
Mais Devagar	1	85	2
Mais Devagar	1	85	2
Mais Devagar	1	86	8
Mais Devagar	1	87	9
Mais Devagar	1	5	11
Mais Devagar	1	5	15
Mais Devagar	1	85	17
Mais Devagar	1	86	20
Mais Devagar	1	86	21
Mais Devagar	2	85	1
Mais Devagar	2	88	5
Mais Devagar	2	86	7
Mais Devagar	2	1	11
Mais Devagar	2	88	12
Mais Devagar	2	1	15
Mais Devagar	2	1	17

Função	Voluntário	Gesto	Tempo
Mais Devagar	3	89	1
Mais Devagar	3	90	3
Mais Devagar	3	5	6
Mais Devagar	3	5	11
Mais Devagar	3	89	13
Mais Devagar	3	5	17
Mais Devagar	4	91	3
Mais Devagar	4	91	6
Mais Devagar	4	91	8
Mais Devagar	4	91	11
Mais Devagar	4	86	14
Mais Devagar	4	86	16
Mais Devagar	4	86	18
Mais Devagar	4	86	20
Mais Devagar	4	86	23
Mais Devagar	5	86	1
Mais Devagar	5	5	6
Mais Devagar	5	86	10
Mais Devagar	5	5	14
Mais Devagar	5	5	20
Mais Devagar	6	92	1
Mais Devagar	6	93	5
Mais Devagar	6	92	8
Mais Devagar	6	93	10
Mais Devagar	6	88	14
Mais Devagar	6	92	16
Mais Devagar	6	93	20
Mais Devagar	7	26	2
Mais Devagar	7	5	5
Mais Devagar	7	3	6
Mais Devagar	7	94	8
Mais Devagar	7	5	10
Mais Devagar	7	26	14
Mais Devagar	7	95	17
Mais Devagar	8	96	2
Mais Devagar	8	88	4
Mais Devagar	8	96	7

Função	Voluntário	Gesto	Tempo
Mais Devagar	8	88	15
Mais Devagar	8	88	16
Mais Devagar	8	96	18
Mais Devagar	8	88	21
Mais Devagar	8	96	23
Mais Devagar	9	3	4
Mais Devagar	9	5	7
Mais Devagar	9	93	9
Mais Devagar	9	93	11
Mais Devagar	9	93	13
Mais Devagar	9	92	17
Mais Devagar	9	93	19
Mais Devagar	9	93	20
Mais Devagar	9	97	25
Mais Devagar	10	85	1
Mais Devagar	10	5	3
Mais Devagar	10	5	4
Mais Devagar	10	85	6
Mais Devagar	10	5	9
Mais Devagar	10	5	12
Mais Devagar	10	85	13
Mais Devagar	10	85	16
Mais Devagar	10	85	18
Mais Devagar	10	5	20
Mais Devagar	10	5	22
Mais Devagar	10	85	26
Mais Devagar	11	85	1
Mais Devagar	11	85	2
Mais Devagar	11	85	5
Mais Devagar	11	43	9
Mais Devagar	11	85	11
Mais Devagar	11	85	14
Mais Devagar	11	5	17
Mais Devagar	11	5	18
Mais Devagar	11	5	20
Mais Devagar	11	5	21
Mais Devagar	11	78	23

Função	Voluntário	Gesto	Tempo
Mais Devagar	11	5	25
Mais Devagar	11	5	26
Mais Devagar	11	5	28
Mais Devagar	11	78	29
Mais Devagar	12	5	2
Mais Devagar	12	5	8
Mais Devagar	12	5	12
Mais Devagar	12	5	17
Mais Devagar	12	2	22
Mais Rápido	1	39	1
Mais Rápido	1	39	5
Mais Rápido	1	36	13
Mais Rápido	1	39	14
Mais Rápido	1	40	22
Mais Rápido	2	30	1
Mais Rápido	2	27	3
Mais Rápido	2	30	4
Mais Rápido	2	30	5
Mais Rápido	2	36	6
Mais Rápido	2	30	8
Mais Rápido	2	31	10
Mais Rápido	2	27	13
Mais Rápido	2	27	16
Mais Rápido	2	27	17
Mais Rápido	2	30	21
Mais Rápido	2	31	23
Mais Rápido	2	40	27
Mais Rápido	2	30	28
Mais Rápido	3	5	2
Mais Rápido	3	41	5
Mais Rápido	3	36	8
Mais Rápido	3	41	10
Mais Rápido	3	36	13
Mais Rápido	3	36	16
Mais Rápido	3	27	19
Mais Rápido	3	27	21
Mais Rápido	4	41	2

Função	Voluntário	Gesto	Tempo
Mais Rápido	4	42	5
Mais Rápido	4	43	8
Mais Rápido	4	44	10
Mais Rápido	4	45	12
Mais Rápido	4	46	14
Mais Rápido	4	47	17
Mais Rápido	4	48	19
Mais Rápido	4	45	23
Mais Rápido	5	39	1
Mais Rápido	5	39	2
Mais Rápido	5	30	4
Mais Rápido	5	39	7
Mais Rápido	5	49	8
Mais Rápido	5	39	9
Mais Rápido	5	49	11
Mais Rápido	5	39	13
Mais Rápido	5	49	14
Mais Rápido	5	39	19
Mais Rápido	5	30	20
Mais Rápido	5	39	22
Mais Rápido	5	49	23
Mais Rápido	6	50	3
Mais Rápido	6	51	5
Mais Rápido	6	51	8
Mais Rápido	6	52	12
Mais Rápido	6	51	16
Mais Rápido	6	51	21
Mais Rápido	7	53	1
Mais Rápido	7	53	3
Mais Rápido	7	53	7
Mais Rápido	7	53	11
Mais Rápido	7	53	14
Mais Rápido	7	22	17
Mais Rápido	7	53	21
Mais Rápido	8	27	1
Mais Rápido	8	27	7
Mais Rápido	8	41	9

Função	Voluntário	Gesto	Tempo
Mais Rápido	8	27	12
Mais Rápido	8	27	14
Mais Rápido	8	27	17
Mais Rápido	8	41	21
Mais Rápido	8	27	25
Mais Rápido	9	27	1
Mais Rápido	9	48	5
Mais Rápido	9	54	9
Mais Rápido	9	48	11
Mais Rápido	9	48	14
Mais Rápido	9	48	16
Mais Rápido	9	48	18
Mais Rápido	9	54	26
Mais Rápido	10	54	1
Mais Rápido	10	48	4
Mais Rápido	10	39	9
Mais Rápido	10	48	10
Mais Rápido	10	55	11
Mais Rápido	10	48	14
Mais Rápido	10	54	16
Mais Rápido	10	55	21
Mais Rápido	10	48	22
Mais Rápido	11	39	3
Mais Rápido	11	48	4
Mais Rápido	11	39	6
Mais Rápido	11	48	11
Mais Rápido	11	41	14
Mais Rápido	11	22	19
Mais Rápido	12	5	2
Mais Rápido	12	5	4
Mais Rápido	12	29	8
Mais Rápido	12	3	10
Mais Rápido	12	22	14
Mais Rápido	12	31	18
Mais Rápido	12	39	21
Mais Rápido	12	5	25
Ok	1	62	1

Função	Voluntário	Gesto	Tempo
Ok	1	57	3
Ok	1	4	5
Ok	1	1	7
Ok	1	5	8
Ok	1	2	10
Ok	1	45	12
Ok	1	3	13
Ok	1	4	15
Ok	1	4	21
Ok	2	64	3
Ok	2	64	7
Ok	2	21	9
Ok	2	21	12
Ok	2	10	14
Ok	2	3	17
Ok	3	3	1
Ok	3	64	2
Ok	3	3	3
Ok	3	11	4
Ok	3	11	6
Ok	3	3	7
Ok	3	11	9
Ok	3	11	10
Ok	3	64	12
Ok	3	3	13
Ok	3	10	14
Ok	3	2	15
Ok	3	2	18
Ok	4	64	1
Ok	4	64	1
Ok	4	65	2
Ok	4	3	3
Ok	4	64	3
Ok	4	32	5
Ok	4	43	7
Ok	4	64	7
Ok	4	65	8

Função	Voluntário	Gesto	Tempo
Ok	4	66	10
Ok	4	3	12
Ok	4	65	14
Ok	4	6	15
Ok	4	2	16
Ok	4	5	19
Ok	4	64	20
Ok	5	43	4
Ok	5	43	7
Ok	5	43	10
Ok	5	43	12
Ok	5	43	16
Ok	5	43	20
Ok	5	43	23
Ok	5	64	24
Ok	6	43	1
Ok	6	5	3
Ok	6	67	4
Ok	6	5	5
Ok	6	3	6
Ok	6	64	7
Ok	6	68	8
Ok	6	70	9
Ok	6	64	12
Ok	6	3	13
Ok	6	71	14
Ok	6	3	16
Ok	6	64	20
Ok	6	72	21
Ok	7	43	1
Ok	7	43	3
Ok	7	73	4
Ok	7	73	5
Ok	7	64	8
Ok	7	5	9
Ok	7	43	11
Ok	7	70	12

Função	Voluntário	Gesto	Tempo
Ok	7	43	14
Ok	7	73	16
Ok	7	73	18
Ok	7	69	21
Ok	8	43	3
Ok	8	5	5
Ok	8	64	6
Ok	8	2	8
Ok	8	43	11
Ok	8	2	12
Ok	8	64	13
Ok	8	3	18
Ok	8	43	19
Ok	8	74	21
Ok	8	64	23
Ok	8	3	23
Ok	9	26	3
Ok	9	43	4
Ok	9	26	4
Ok	9	43	11
Ok	9	5	13
Ok	9	75	17
Ok	9	43	20
Ok	9	64	21
Ok	9	75	22
Ok	9	5	23
Ok	10	43	7
Ok	10	5	11
Ok	10	5	15
Ok	10	5	20
Ok	10	5	25
Ok	11	43	1
Ok	11	15	6
Ok	11	5	11
Ok	11	5	23
Ok	11	3	29
Ok	12	5	1

Função	Voluntário	Gesto	Tempo
Ok	12	64	2
Ok	12	64	6
Ok	12	3	8
Ok	12	29	9
Ok	12	64	13
Ok	12	3	20
Ok	12	64	22
Ok	12	43	23
Ok	12	64	25
Siga-me	1	36	1
Siga-me	1	36	3
Siga-me	1	36	11
Siga-me	1	36	14
Siga-me	1	36	16
Siga-me	1	37	20
Siga-me	1	37	23
Siga-me	1	37	27
Siga-me	2	27	1
Siga-me	2	22	3
Siga-me	2	27	7
Siga-me	2	56	8
Siga-me	2	56	11
Siga-me	2	27	12
Siga-me	2	27	14
Siga-me	2	37	16
Siga-me	2	56	18
Siga-me	2	56	20
Siga-me	2	56	22
Siga-me	2	27	24
Siga-me	2	22	25
Siga-me	2	56	28
Siga-me	3	27	1
Siga-me	3	27	3
Siga-me	3	22	5
Siga-me	3	22	9
Siga-me	3	57	12
Siga-me	3	56	16

Função	Voluntário	Gesto	Tempo
Siga-me	3	58	17
Siga-me	3	22	21
Siga-me	3	22	23
Siga-me	4	59	1
Siga-me	4	60	3
Siga-me	4	27	6
Siga-me	4	27	7
Siga-me	4	59	10
Siga-me	4	59	13
Siga-me	4	59	14
Siga-me	4	22	16
Siga-me	4	59	17
Siga-me	4	59	20
Siga-me	5	27	2
Siga-me	5	56	4
Siga-me	5	50	9
Siga-me	5	50	12
Siga-me	5	50	14
Siga-me	5	27	16
Siga-me	5	27	18
Siga-me	5	56	24
Siga-me	6	27	1
Siga-me	6	27	2
Siga-me	6	56	5
Siga-me	6	27	8
Siga-me	6	27	9
Siga-me	6	22	12
Siga-me	6	50	12
Siga-me	6	30	15
Siga-me	6	36	20
Siga-me	6	43	23
Siga-me	7	31	3
Siga-me	7	36	8
Siga-me	7	27	12
Siga-me	7	27	20
Siga-me	7	36	25
Siga-me	8	36	1

Função	Voluntário	Gesto	Tempo
Siga-me	8	54	4
Siga-me	8	54	7
Siga-me	8	53	9
Siga-me	8	53	11
Siga-me	8	36	12
Siga-me	8	54	15
Siga-me	8	36	19
Siga-me	8	36	22
Siga-me	9	27	1
Siga-me	9	56	3
Siga-me	9	27	7
Siga-me	9	27	9
Siga-me	9	58	11
Siga-me	9	27	13
Siga-me	9	60	15
Siga-me	9	50	16
Siga-me	9	60	19
Siga-me	9	60	21
Siga-me	9	27	23
Siga-me	9	27	25
Siga-me	10	27	1
Siga-me	10	36	7
Siga-me	10	22	10
Siga-me	10	36	17
Siga-me	10	31	20
Siga-me	11	27	2
Siga-me	11	62	5
Siga-me	11	61	11
Siga-me	11	62	13
Siga-me	11	62	17
Siga-me	11	27	19
Siga-me	11	27	23
Siga-me	12	22	1
Siga-me	12	27	2
Siga-me	12	27	3
Siga-me	12	62	8
Siga-me	12	4	13

Função	Voluntário	Gesto	Tempo
Siga-me	12	27	19
Siga-me	12	27	22

Apêndice D – Tabelas com as taxas de ocorrências dos gestos

Tabela 7 – Taxas de Ocorrência Geral (TOG) e pesos obtidos para cada gesto em cada função. As linhas estão ordenadas respectivamente pela função e pelo peso

Função	Gesto	TOG	Peso
Abortar	2	0,375	1
Abortar	5	0,1518	0,9474
Abortar	3	0,1339	0,8947
Abortar	8	0,0357	0,8421
Abortar	4	0,0357	0,8421
Abortar	17	0,0357	0,8421
Abortar	19	0,0357	0,8421
Abortar	18	0,0357	0,8421
Abortar	16	0,0268	0,5789
Abortar	1	0,0179	0,5263
Abortar	6	0,0179	0,5263
Abortar	12	0,0179	0,5263
Abortar	11	0,0179	0,5263
Abortar	10	0,0179	0,5263
Abortar	13	0,0089	0,2632
Abortar	9	0,0089	0,2632
Abortar	15	0,0089	0,2632
Abortar	14	0,0089	0,2632
Abortar	7	0,0089	0,2632
Apontar	76	0,413	1
Apontar	78	0,163	0,9375
Apontar	27	0,0761	0,875
Apontar	79	0,0652	0,8125
Apontar	77	0,0543	0,75
Apontar	81	0,0435	0,6875
Apontar	82	0,0435	0,6875
Apontar	31	0,0326	0,5625
Apontar	5	0,0326	0,5625
Apontar	84	0,0109	0,4375
Apontar	80	0,0109	0,4375

Função	Gesto	TOG	Peso
Apontar	3	0,0109	0,4375
Apontar	26	0,0109	0,4375
Apontar	59	0,0109	0,4375
Apontar	37	0,0109	0,4375
Apontar	83	0,0109	0,4375
Atenção	28	0,215	1
Atenção	26	0,1121	0,9583
Atenção	4	0,0841	0,9167
Atenção	27	0,0748	0,875
Atenção	21	0,0654	0,8333
Atenção	33	0,0654	0,8333
Atenção	29	0,0561	0,75
Atenção	22	0,0561	0,75
Atenção	37	0,0374	0,6667
Atenção	31	0,0374	0,6667
Atenção	32	0,0374	0,6667
Atenção	38	0,0187	0,5417
Atenção	36	0,0187	0,5417
Atenção	25	0,0187	0,5417
Atenção	5	0,0187	0,5417
Atenção	24	0,0093	0,375
Atenção	35	0,0093	0,375
Atenção	2	0,0093	0,375
Atenção	23	0,0093	0,375
Atenção	1	0,0093	0,375
Atenção	30	0,0093	0,375
Atenção	42	0,0093	0,375
Atenção	34	0,0093	0,375
Atenção	11	0,0093	0,375
Mais Devagar	5	0,2828	1
Mais Devagar	85	0,1515	0,95
Mais Devagar	86	0,1111	0,9
Mais Devagar	93	0,0808	0,85
Mais Devagar	88	0,0707	0,8
Mais Devagar	92	0,0404	0,75
Mais Devagar	91	0,0404	0,75
Mais Devagar	96	0,0404	0,75

Função	Gesto	TOG	Peso
Mais Devagar	1	0,0303	0,6
Mais Devagar	3	0,0202	0,55
Mais Devagar	26	0,0202	0,55
Mais Devagar	78	0,0202	0,55
Mais Devagar	89	0,0202	0,55
Mais Devagar	97	0,0101	0,35
Mais Devagar	95	0,0101	0,35
Mais Devagar	94	0,0101	0,35
Mais Devagar	2	0,0101	0,35
Mais Devagar	90	0,0101	0,35
Mais Devagar	87	0,0101	0,35
Mais Devagar	43	0,0101	0,35
Mais Rápido	39	0,1386	1
Mais Rápido	27	0,1287	0,96
Mais Rápido	48	0,1188	0,92
Mais Rápido	30	0,0792	0,88
Mais Rápido	53	0,0594	0,84
Mais Rápido	41	0,0594	0,84
Mais Rápido	36	0,0495	0,76
Mais Rápido	49	0,0396	0,72
Mais Rápido	54	0,0396	0,72
Mais Rápido	5	0,0396	0,72
Mais Rápido	51	0,0396	0,72
Mais Rápido	22	0,0297	0,56
Mais Rápido	31	0,0297	0,56
Mais Rápido	40	0,0198	0,48
Mais Rápido	45	0,0198	0,48
Mais Rápido	55	0,0198	0,48
Mais Rápido	50	0,0099	0,36
Mais Rápido	44	0,0099	0,36
Mais Rápido	52	0,0099	0,36
Mais Rápido	47	0,0099	0,36
Mais Rápido	46	0,0099	0,36
Mais Rápido	43	0,0099	0,36
Mais Rápido	29	0,0099	0,36
Mais Rápido	3	0,0099	0,36
Mais Rápido	42	0,0099	0,36

Função	Gesto	TOG	Peso
Ok	64	0,1901	1
Ok	43	0,1818	0,9655
Ok	3	0,1322	0,931
Ok	5	0,124	0,8966
Ok	2	0,0496	0,8621
Ok	11	0,0331	0,8276
Ok	73	0,0331	0,8276
Ok	4	0,0248	0,7586
Ok	65	0,0248	0,7586
Ok	10	0,0165	0,6897
Ok	26	0,0165	0,6897
Ok	70	0,0165	0,6897
Ok	75	0,0165	0,6897
Ok	21	0,0165	0,6897
Ok	66	0,0083	0,5172
Ok	57	0,0083	0,5172
Ok	29	0,0083	0,5172
Ok	1	0,0083	0,5172
Ok	6	0,0083	0,5172
Ok	74	0,0083	0,5172
Ok	72	0,0083	0,5172
Ok	71	0,0083	0,5172
Ok	68	0,0083	0,5172
Ok	67	0,0083	0,5172
Ok	62	0,0083	0,5172
Ok	15	0,0083	0,5172
Ok	45	0,0083	0,5172
Ok	32	0,0083	0,5172
Ok	69	0,0083	0,5172
Siga-me	27	0,3077	1
Siga-me	36	0,1346	0,9444
Siga-me	56	0,1058	0,8889
Siga-me	22	0,0962	0,8333
Siga-me	59	0,0577	0,7778
Siga-me	50	0,0481	0,7222
Siga-me	37	0,0385	0,6667
Siga-me	62	0,0385	0,6667

Função	Gesto	TOG	Peso
Siga-me	60	0,0385	0,6667
Siga-me	54	0,0288	0,5
Siga-me	53	0,0192	0,4444
Siga-me	31	0,0192	0,4444
Siga-me	58	0,0192	0,4444
Siga-me	4	0,0096	0,2778
Siga-me	30	0,0096	0,2778
Siga-me	57	0,0096	0,2778
Siga-me	61	0,0096	0,2778
Siga-me	43	0,0096	0,2778

Tabela 8 – Taxas de Primeira Ocorrência (TPO) e pesos obtidos para cada gesto em cada função. As linhas estão ordenadas respectivamente pela função e pelo peso

Função	Gesto	TPO	Peso
Abortar	2	0,2391	1
Abortar	3	0,1304	0,9474
Abortar	5	0,1304	0,9474
Abortar	4	0,0435	0,8421
Abortar	17	0,0435	0,8421
Abortar	18	0,0435	0,8421
Abortar	16	0,0435	0,8421
Abortar	12	0,0435	0,8421
Abortar	8	0,0435	0,8421
Abortar	11	0,0435	0,8421
Abortar	19	0,0217	0,4737
Abortar	1	0,0217	0,4737
Abortar	9	0,0217	0,4737
Abortar	10	0,0217	0,4737
Abortar	14	0,0217	0,4737
Abortar	13	0,0217	0,4737
Abortar	7	0,0217	0,4737
Abortar	15	0,0217	0,4737
Abortar	6	0,0217	0,4737

Função	Gesto	TPO	Peso
Apontar	76	0,2105	1
Apontar	78	0,1316	0,9375
Apontar	79	0,1053	0,875
Apontar	77	0,1053	0,875
Apontar	27	0,0789	0,75
Apontar	81	0,0526	0,6875
Apontar	31	0,0526	0,6875
Apontar	5	0,0526	0,6875
Apontar	37	0,0263	0,5
Apontar	83	0,0263	0,5
Apontar	3	0,0263	0,5
Apontar	84	0,0263	0,5
Apontar	59	0,0263	0,5
Apontar	82	0,0263	0,5
Apontar	26	0,0263	0,5
Apontar	80	0,0263	0,5
Atenção	28	0,1538	1
Atenção	26	0,0962	0,9583
Atenção	22	0,0962	0,9583
Atenção	27	0,0769	0,875
Atenção	4	0,0769	0,875
Atenção	21	0,0769	0,875
Atenção	25	0,0385	0,75
Atenção	37	0,0385	0,75
Atenção	29	0,0385	0,75
Atenção	31	0,0385	0,75
Atenção	35	0,0192	0,5833
Atenção	32	0,0192	0,5833
Atenção	11	0,0192	0,5833
Atenção	5	0,0192	0,5833
Atenção	30	0,0192	0,5833
Atenção	2	0,0192	0,5833
Atenção	23	0,0192	0,5833
Atenção	38	0,0192	0,5833
Atenção	33	0,0192	0,5833
Atenção	42	0,0192	0,5833
Atenção	36	0,0192	0,5833

Função	Gesto	TPO	Peso
Atenção	34	0,0192	0,5833
Atenção	24	0,0192	0,5833
Atenção	1	0,0192	0,5833
Mais Devagar	5	0,2105	1
Mais Devagar	86	0,1053	0,95
Mais Devagar	85	0,1053	0,95
Mais Devagar	88	0,0789	0,85
Mais Devagar	93	0,0526	0,8
Mais Devagar	3	0,0526	0,8
Mais Devagar	92	0,0526	0,8
Mais Devagar	87	0,0263	0,65
Mais Devagar	91	0,0263	0,65
Mais Devagar	1	0,0263	0,65
Mais Devagar	97	0,0263	0,65
Mais Devagar	2	0,0263	0,65
Mais Devagar	43	0,0263	0,65
Mais Devagar	89	0,0263	0,65
Mais Devagar	90	0,0263	0,65
Mais Devagar	94	0,0263	0,65
Mais Devagar	96	0,0263	0,65
Mais Devagar	95	0,0263	0,65
Mais Devagar	78	0,0263	0,65
Mais Devagar	26	0,0263	0,65
Mais Rápido	39	0,1064	1
Mais Rápido	41	0,0851	0,96
Mais Rápido	48	0,0851	0,96
Mais Rápido	27	0,0851	0,96
Mais Rápido	36	0,0638	0,84
Mais Rápido	22	0,0638	0,84
Mais Rápido	30	0,0426	0,76
Mais Rápido	40	0,0426	0,76
Mais Rápido	31	0,0426	0,76
Mais Rápido	5	0,0426	0,76
Mais Rápido	54	0,0426	0,76
Mais Rápido	46	0,0213	0,56
Mais Rápido	43	0,0213	0,56
Mais Rápido	52	0,0213	0,56

Função	Gesto	TPO	Peso
Mais Rápido	49	0,0213	0,56
Mais Rápido	42	0,0213	0,56
Mais Rápido	3	0,0213	0,56
Mais Rápido	47	0,0213	0,56
Mais Rápido	45	0,0213	0,56
Mais Rápido	29	0,0213	0,56
Mais Rápido	44	0,0213	0,56
Mais Rápido	53	0,0213	0,56
Mais Rápido	51	0,0213	0,56
Mais Rápido	55	0,0213	0,56
Mais Rápido	50	0,0213	0,56
Ok	64	0,1385	1
Ok	43	0,1385	1
Ok	5	0,1385	1
Ok	3	0,1231	0,8966
Ok	2	0,0615	0,8621
Ok	10	0,0308	0,8276
Ok	70	0,0308	0,8276
Ok	1	0,0154	0,7586
Ok	57	0,0154	0,7586
Ok	45	0,0154	0,7586
Ok	32	0,0154	0,7586
Ok	29	0,0154	0,7586
Ok	26	0,0154	0,7586
Ok	21	0,0154	0,7586
Ok	15	0,0154	0,7586
Ok	6	0,0154	0,7586
Ok	4	0,0154	0,7586
Ok	11	0,0154	0,7586
Ok	74	0,0154	0,7586
Ok	65	0,0154	0,7586
Ok	73	0,0154	0,7586
Ok	66	0,0154	0,7586
Ok	67	0,0154	0,7586
Ok	68	0,0154	0,7586
Ok	69	0,0154	0,7586
Ok	71	0,0154	0,7586

Função	Gesto	TPO	Peso
Ok	72	0,0154	0,7586
Ok	75	0,0154	0,7586
Ok	62	0,0154	0,7586
Siga-me	27	0,2128	1
Siga-me	22	0,1277	0,9444
Siga-me	56	0,1064	0,8889
Siga-me	36	0,1064	0,8889
Siga-me	50	0,0638	0,7778
Siga-me	37	0,0426	0,7222
Siga-me	60	0,0426	0,7222
Siga-me	31	0,0426	0,7222
Siga-me	62	0,0426	0,7222
Siga-me	58	0,0426	0,7222
Siga-me	53	0,0213	0,4444
Siga-me	54	0,0213	0,4444
Siga-me	57	0,0213	0,4444
Siga-me	30	0,0213	0,4444
Siga-me	59	0,0213	0,4444
Siga-me	4	0,0213	0,4444
Siga-me	61	0,0213	0,4444
Siga-me	43	0,0213	0,4444

Tabela 9 – Taxas de Ocorrência por Voluntário (TOV) e pesos obtidos para cada gesto em cada função. As linhas estão ordenadas respectivamente pela função e pelo peso

Função	Gesto	TOV	Peso
Abortar	2	0,9167	1
Abortar	3	0,5	0,9474
Abortar	5	0,5	0,9474
Abortar	4	0,1667	0,8421
Abortar	17	0,1667	0,8421
Abortar	18	0,1667	0,8421
Abortar	16	0,1667	0,8421
Abortar	12	0,1667	0,8421

Função	Gesto	TOV	Peso
Abortar	11	0,1667	0,8421
Abortar	8	0,1667	0,8421
Abortar	7	0,0833	0,4737
Abortar	19	0,0833	0,4737
Abortar	9	0,0833	0,4737
Abortar	6	0,0833	0,4737
Abortar	10	0,0833	0,4737
Abortar	14	0,0833	0,4737
Abortar	13	0,0833	0,4737
Abortar	15	0,0833	0,4737
Abortar	1	0,0833	0,4737
Apontar	76	0,6667	1
Apontar	78	0,4167	0,9375
Apontar	79	0,3333	0,875
Apontar	77	0,3333	0,875
Apontar	27	0,25	0,75
Apontar	81	0,1667	0,6875
Apontar	31	0,1667	0,6875
Apontar	5	0,1667	0,6875
Apontar	37	0,0833	0,5
Apontar	83	0,0833	0,5
Apontar	3	0,0833	0,5
Apontar	84	0,0833	0,5
Apontar	59	0,0833	0,5
Apontar	82	0,0833	0,5
Apontar	26	0,0833	0,5
Apontar	80	0,0833	0,5
Atenção	28	0,6667	1
Atenção	22	0,4167	0,9583
Atenção	26	0,4167	0,9583
Atenção	21	0,3333	0,875
Atenção	27	0,3333	0,875
Atenção	4	0,3333	0,875
Atenção	29	0,1667	0,75
Atenção	37	0,1667	0,75
Atenção	31	0,1667	0,75
Atenção	25	0,1667	0,75

Função	Gesto	TOV	Peso
Atenção	42	0,0833	0,5833
Atenção	1	0,0833	0,5833
Atenção	24	0,0833	0,5833
Atenção	35	0,0833	0,5833
Atenção	23	0,0833	0,5833
Atenção	11	0,0833	0,5833
Atenção	5	0,0833	0,5833
Atenção	30	0,0833	0,5833
Atenção	2	0,0833	0,5833
Atenção	32	0,0833	0,5833
Atenção	33	0,0833	0,5833
Atenção	36	0,0833	0,5833
Atenção	34	0,0833	0,5833
Atenção	38	0,0833	0,5833
Mais Devagar	5	0,6667	1
Mais Devagar	86	0,3333	0,95
Mais Devagar	85	0,3333	0,95
Mais Devagar	88	0,25	0,85
Mais Devagar	93	0,1667	0,8
Mais Devagar	3	0,1667	0,8
Mais Devagar	92	0,1667	0,8
Mais Devagar	87	0,0833	0,65
Mais Devagar	91	0,0833	0,65
Mais Devagar	1	0,0833	0,65
Mais Devagar	97	0,0833	0,65
Mais Devagar	2	0,0833	0,65
Mais Devagar	43	0,0833	0,65
Mais Devagar	89	0,0833	0,65
Mais Devagar	90	0,0833	0,65
Mais Devagar	94	0,0833	0,65
Mais Devagar	96	0,0833	0,65
Mais Devagar	95	0,0833	0,65
Mais Devagar	78	0,0833	0,65
Mais Devagar	26	0,0833	0,65
Mais Rápido	39	0,4167	1
Mais Rápido	41	0,3333	0,96
Mais Rápido	48	0,3333	0,96

Função	Gesto	TOV	Peso
Mais Rápido	27	0,3333	0,96
Mais Rápido	22	0,25	0,84
Mais Rápido	36	0,25	0,84
Mais Rápido	5	0,1667	0,76
Mais Rápido	54	0,1667	0,76
Mais Rápido	31	0,1667	0,76
Mais Rápido	40	0,1667	0,76
Mais Rápido	30	0,1667	0,76
Mais Rápido	29	0,0833	0,56
Mais Rápido	46	0,0833	0,56
Mais Rápido	45	0,0833	0,56
Mais Rápido	42	0,0833	0,56
Mais Rápido	44	0,0833	0,56
Mais Rápido	3	0,0833	0,56
Mais Rápido	55	0,0833	0,56
Mais Rápido	52	0,0833	0,56
Mais Rápido	53	0,0833	0,56
Mais Rápido	49	0,0833	0,56
Mais Rápido	47	0,0833	0,56
Mais Rápido	51	0,0833	0,56
Mais Rápido	50	0,0833	0,56
Mais Rápido	43	0,0833	0,56
Ok	64	0,75	1
Ok	43	0,75	1
Ok	5	0,75	1
Ok	3	0,6667	0,8966
Ok	2	0,3333	0,8621
Ok	10	0,1667	0,8276
Ok	70	0,1667	0,8276
Ok	1	0,0833	0,7586
Ok	57	0,0833	0,7586
Ok	45	0,0833	0,7586
Ok	32	0,0833	0,7586
Ok	29	0,0833	0,7586
Ok	26	0,0833	0,7586
Ok	21	0,0833	0,7586
Ok	15	0,0833	0,7586

Função	Gesto	TOV	Peso
Ok	6	0,0833	0,7586
Ok	4	0,0833	0,7586
Ok	11	0,0833	0,7586
Ok	74	0,0833	0,7586
Ok	65	0,0833	0,7586
Ok	73	0,0833	0,7586
Ok	66	0,0833	0,7586
Ok	67	0,0833	0,7586
Ok	68	0,0833	0,7586
Ok	69	0,0833	0,7586
Ok	71	0,0833	0,7586
Ok	72	0,0833	0,7586
Ok	75	0,0833	0,7586
Ok	62	0,0833	0,7586
Siga-me	27	0,8333	1
Siga-me	22	0,5	0,9444
Siga-me	56	0,4167	0,8889
Siga-me	36	0,4167	0,8889
Siga-me	50	0,25	0,7778
Siga-me	60	0,1667	0,7222
Siga-me	37	0,1667	0,7222
Siga-me	31	0,1667	0,7222
Siga-me	62	0,1667	0,7222
Siga-me	58	0,1667	0,7222
Siga-me	53	0,0833	0,4444
Siga-me	54	0,0833	0,4444
Siga-me	57	0,0833	0,4444
Siga-me	30	0,0833	0,4444
Siga-me	59	0,0833	0,4444
Siga-me	4	0,0833	0,4444
Siga-me	61	0,0833	0,4444
Siga-me	43	0,0833	0,4444

Tabela 10 – Taxas de Ocorrência por Tempo (TOT) e pesos obtidos para cada gesto em cada função. As linhas estão ordenadas respectivamente pela função e pelo peso

Função	Gesto	TOT	Peso
Abortar	2	5,1377	1
Abortar	5	1,7413	0,9474
Abortar	3	0,8221	0,8947
Abortar	8	0,6223	0,8421
Abortar	4	0,6	0,7895
Abortar	18	0,5971	0,7368
Abortar	19	0,5789	0,6842
Abortar	13	0,4667	0,6316
Abortar	1	0,2294	0,5789
Abortar	16	0,2164	0,5263
Abortar	10	0,1724	0,4737
Abortar	17	0,1715	0,4211
Abortar	6	0,1412	0,3684
Abortar	9	0,1207	0,3158
Abortar	12	0,0995	0,2632
Abortar	15	0,0976	0,2105
Abortar	11	0,0937	0,1579
Abortar	14	0,0857	0,1053
Abortar	7	0,0059	0,0526
Apontar	76	4,5466	1
Apontar	78	2,2103	0,9375
Apontar	27	0,9596	0,875
Apontar	77	0,7792	0,8125
Apontar	81	0,7583	0,75
Apontar	79	0,602	0,6875
Apontar	31	0,541	0,625
Apontar	82	0,4706	0,5625
Apontar	80	0,3286	0,5
Apontar	3	0,2571	0,4375
Apontar	5	0,209	0,375
Apontar	83	0,1345	0,3125
Apontar	37	0,0833	0,25
Apontar	84	0,0672	0,1875

Função	Gesto	TOT	Peso
Apontar	26	0,0385	0,125
Apontar	59	0,0143	0,0625
Atenção	28	2,9376	1
Atenção	21	1,3309	0,9583
Atenção	4	1,1381	0,9167
Atenção	26	1,0455	0,875
Atenção	22	0,8609	0,8333
Atenção	27	0,8039	0,7917
Atenção	37	0,5681	0,75
Atenção	29	0,5495	0,7083
Atenção	33	0,4568	0,6667
Atenção	32	0,4198	0,625
Atenção	31	0,2752	0,5833
Atenção	34	0,2308	0,5417
Atenção	35	0,2154	0,5
Atenção	38	0,1957	0,4583
Atenção	23	0,181	0,4167
Atenção	42	0,1338	0,375
Atenção	36	0,1235	0,3333
Atenção	24	0,1121	0,2917
Atenção	30	0,1106	0,25
Atenção	25	0,091	0,2083
Atenção	11	0,0821	0,1667
Atenção	5	0,0632	0,125
Atenção	1	0,0597	0,0833
Atenção	2	0,0151	0,0417
Mais Devagar	5	3,2566	1
Mais Devagar	85	1,8303	0,95
Mais Devagar	86	1,2455	0,9
Mais Devagar	93	0,9157	0,85
Mais Devagar	88	0,8886	0,8
Mais Devagar	91	0,701	0,75
Mais Devagar	92	0,6031	0,7
Mais Devagar	96	0,5349	0,65
Mais Devagar	3	0,3893	0,6
Mais Devagar	89	0,386	0,55
Mais Devagar	26	0,3125	0,5

Função	Gesto	TOT	Peso
Mais Devagar	90	0,2632	0,45
Mais Devagar	1	0,1897	0,4
Mais Devagar	94	0,1562	0,35
Mais Devagar	87	0,1398	0,3
Mais Devagar	43	0,1045	0,25
Mais Devagar	78	0,0398	0,2
Mais Devagar	2	0,0185	0,15
Mais Devagar	95	0,0156	0,1
Mais Devagar	97	0,0092	0,05
Mais Rápido	39	2,0948	1
Mais Rápido	48	1,4803	0,96
Mais Rápido	27	1,3563	0,92
Mais Rápido	53	0,9375	0,88
Mais Rápido	41	0,8721	0,84
Mais Rápido	5	0,6873	0,8
Mais Rápido	30	0,6286	0,76
Mais Rápido	36	0,623	0,72
Mais Rápido	51	0,5672	0,68
Mais Rápido	54	0,4567	0,64
Mais Rápido	50	0,2836	0,6
Mais Rápido	49	0,2516	0,56
Mais Rápido	22	0,1916	0,52
Mais Rápido	31	0,1871	0,48
Mais Rápido	42	0,1792	0,44
Mais Rápido	29	0,1698	0,4
Mais Rápido	43	0,1509	0,36
Mais Rápido	3	0,1509	0,36
Mais Rápido	52	0,1493	0,28
Mais Rápido	55	0,1414	0,24
Mais Rápido	44	0,1321	0,2
Mais Rápido	45	0,1226	0,16
Mais Rápido	46	0,0943	0,12
Mais Rápido	47	0,066	0,08
Mais Rápido	40	0,0256	0,04
Ok	43	3,0234	1
Ok	64	2,135	0,9655
Ok	5	1,8458	0,931

Função	Gesto	TOT	Peso
Ok	3	1,0593	0,8966
Ok	26	0,402	0,8621
Ok	2	0,3817	0,8276
Ok	11	0,3534	0,7931
Ok	21	0,3261	0,7586
Ok	73	0,3169	0,7241
Ok	15	0,3	0,6897
Ok	65	0,2021	0,6552
Ok	4	0,2	0,6207
Ok	62	0,168	0,5862
Ok	57	0,152	0,5517
Ok	70	0,1473	0,5172
Ok	29	0,1298	0,4828
Ok	10	0,1246	0,4483
Ok	1	0,12	0,4138
Ok	67	0,1065	0,3793
Ok	75	0,0882	0,3448
Ok	32	0,0829	0,3103
Ok	68	0,0828	0,2759
Ok	45	0,08	0,2414
Ok	66	0,057	0,2069
Ok	71	0,0473	0,1724
Ok	6	0,0311	0,1379
Ok	74	0,0238	0,1034
Ok	69	0,007	0,069
Ok	72	0,0059	0,0345
Siga-me	27	4,1841	1
Siga-me	36	1,9232	0,9444
Siga-me	22	1,0962	0,8889
Siga-me	56	0,9285	0,8333
Siga-me	62	0,6356	0,7778
Siga-me	50	0,5534	0,7222
Siga-me	59	0,4951	0,6667
Siga-me	54	0,4019	0,6111
Siga-me	31	0,391	0,5556
Siga-me	60	0,3291	0,5
Siga-me	53	0,243	0,4444

Função	Gesto	TOT	Peso
Siga-me	37	0,1944	0,3889
Siga-me	61	0,1667	0,3333
Siga-me	58	0,1649	0,2778
Siga-me	57	0,1101	0,2222
Siga-me	4	0,1075	0,1667
Siga-me	30	0,0677	0,1111
Siga-me	43	0,0075	0,0556

Apêndice E – Pesos dos gestos em cada função

Tabela 11 – Pesos de cada gesto em cada função. As linhas estão ordenados respectivamente pela função e pelo Peso Geral (PG)

Função	Gesto	Peso TOG	Peso TPO	Peso TOV	Peso TOT	PG
Abortar	2	1	1	1	1	1
Abortar	5	0,9474	0,9474	0,9474	0,9474	0,9474
Abortar	3	0,8947	0,9474	0,9474	0,8947	0,92105
Abortar	8	0,8421	0,8421	0,8421	0,8421	0,8421
Abortar	4	0,8421	0,8421	0,8421	0,7895	0,82895
Abortar	18	0,8421	0,8421	0,8421	0,7368	0,815775
Abortar	17	0,8421	0,8421	0,8421	0,4211	0,73685
Abortar	16	0,5789	0,8421	0,8421	0,5263	0,69735
Abortar	19	0,8421	0,4737	0,4737	0,6842	0,618425
Abortar	12	0,5263	0,8421	0,8421	0,2632	0,618425
Abortar	11	0,5263	0,8421	0,8421	0,1579	0,5921
Abortar	1	0,5263	0,4737	0,4737	0,5789	0,51315
Abortar	10	0,5263	0,4737	0,4737	0,4737	0,48685
Abortar	13	0,2632	0,4737	0,4737	0,6316	0,46055
Abortar	6	0,5263	0,4737	0,4737	0,3684	0,460525
Abortar	9	0,2632	0,4737	0,4737	0,3158	0,3816
Abortar	15	0,2632	0,4737	0,4737	0,2105	0,355275
Abortar	14	0,2632	0,4737	0,4737	0,1053	0,328975
Abortar	7	0,2632	0,4737	0,4737	0,0526	0,3158
Apontar	76	1	1	1	1	1
Apontar	78	0,9375	0,9375	0,9375	0,9375	0,9375
Apontar	77	0,75	0,875	0,875	0,8125	0,828125
Apontar	79	0,8125	0,875	0,875	0,6875	0,8125
Apontar	27	0,875	0,75	0,75	0,875	0,8125
Apontar	81	0,6875	0,6875	0,6875	0,75	0,703125
Apontar	31	0,5625	0,6875	0,6875	0,625	0,640625
Apontar	5	0,5625	0,6875	0,6875	0,375	0,578125
Apontar	82	0,6875	0,5	0,5	0,5625	0,5625
Apontar	80	0,4375	0,5	0,5	0,5	0,484375
Apontar	3	0,4375	0,5	0,5	0,4375	0,46875
Apontar	83	0,4375	0,5	0,5	0,3125	0,4375

Função	Gesto	Peso TOG	Peso TPO	Peso TOV	Peso TOT	PG
Apontar	37	0,4375	0,5	0,5	0,25	0,421875
Apontar	84	0,4375	0,5	0,5	0,1875	0,40625
Apontar	26	0,4375	0,5	0,5	0,125	0,390625
Apontar	59	0,4375	0,5	0,5	0,0625	0,375
Atenção	28	1	1	1	1	1
Atenção	26	0,9583	0,9583	0,9583	0,875	0,937475
Atenção	4	0,9167	0,875	0,875	0,9167	0,89585
Atenção	21	0,8333	0,875	0,875	0,9583	0,8854
Atenção	22	0,75	0,9583	0,9583	0,8333	0,874975
Atenção	27	0,875	0,875	0,875	0,7917	0,854175
Atenção	29	0,75	0,75	0,75	0,7083	0,739575
Atenção	37	0,6667	0,75	0,75	0,75	0,729175
Atenção	31	0,6667	0,75	0,75	0,5833	0,6875
Atenção	33	0,8333	0,5833	0,5833	0,6667	0,66665
Atenção	32	0,6667	0,5833	0,5833	0,625	0,614575
Atenção	25	0,5417	0,75	0,75	0,2083	0,5625
Atenção	38	0,5417	0,5833	0,5833	0,4583	0,54165
Atenção	34	0,375	0,5833	0,5833	0,5417	0,520825
Atenção	35	0,375	0,5833	0,5833	0,5	0,5104
Atenção	36	0,5417	0,5833	0,5833	0,3333	0,5104
Atenção	23	0,375	0,5833	0,5833	0,4167	0,489575
Atenção	42	0,375	0,5833	0,5833	0,375	0,47915
Atenção	24	0,375	0,5833	0,5833	0,2917	0,458325
Atenção	5	0,5417	0,5833	0,5833	0,125	0,458325
Atenção	30	0,375	0,5833	0,5833	0,25	0,4479
Atenção	11	0,375	0,5833	0,5833	0,1667	0,427075
Atenção	1	0,375	0,5833	0,5833	0,0833	0,406225
Atenção	2	0,375	0,5833	0,5833	0,0417	0,395825
Mais Devagar	5	1	1	1	1	1
Mais Devagar	85	0,95	0,95	0,95	0,95	0,95
Mais Devagar	86	0,9	0,95	0,95	0,9	0,925
Mais Devagar	93	0,85	0,8	0,8	0,85	0,825
Mais Devagar	88	0,8	0,85	0,85	0,8	0,825
Mais Devagar	92	0,75	0,8	0,8	0,7	0,7625
Mais Devagar	91	0,75	0,65	0,65	0,75	0,7
Mais Devagar	3	0,55	0,8	0,8	0,6	0,6875
Mais Devagar	96	0,75	0,65	0,65	0,65	0,675

Função	Gesto	Peso TOG	Peso TPO	Peso TOV	Peso TOT	PG
Mais Devagar	89	0,55	0,65	0,65	0,55	0,6
Mais Devagar	26	0,55	0,65	0,65	0,5	0,5875
Mais Devagar	1	0,6	0,65	0,65	0,4	0,575
Mais Devagar	90	0,35	0,65	0,65	0,45	0,525
Mais Devagar	78	0,55	0,65	0,65	0,2	0,5125
Mais Devagar	94	0,35	0,65	0,65	0,35	0,5
Mais Devagar	87	0,35	0,65	0,65	0,3	0,4875
Mais Devagar	43	0,35	0,65	0,65	0,25	0,475
Mais Devagar	2	0,35	0,65	0,65	0,15	0,45
Mais Devagar	95	0,35	0,65	0,65	0,1	0,4375
Mais Devagar	97	0,35	0,65	0,65	0,05	0,425
Mais Rápido	39	1	1	1	1	1
Mais Rápido	48	0,92	0,96	0,96	0,96	0,95
Mais Rápido	27	0,96	0,96	0,96	0,92	0,95
Mais Rápido	41	0,84	0,96	0,96	0,84	0,9
Mais Rápido	36	0,76	0,84	0,84	0,72	0,79
Mais Rápido	30	0,88	0,76	0,76	0,76	0,79
Mais Rápido	5	0,72	0,76	0,76	0,8	0,76
Mais Rápido	54	0,72	0,76	0,76	0,64	0,72
Mais Rápido	53	0,84	0,56	0,56	0,88	0,71
Mais Rápido	22	0,56	0,84	0,84	0,52	0,69
Mais Rápido	31	0,56	0,76	0,76	0,48	0,64
Mais Rápido	51	0,72	0,56	0,56	0,68	0,63
Mais Rápido	49	0,72	0,56	0,56	0,56	0,6
Mais Rápido	50	0,36	0,56	0,56	0,6	0,52
Mais Rápido	40	0,48	0,76	0,76	0,04	0,51
Mais Rápido	42	0,36	0,56	0,56	0,44	0,48
Mais Rápido	29	0,36	0,56	0,56	0,4	0,47
Mais Rápido	55	0,48	0,56	0,56	0,24	0,46
Mais Rápido	3	0,36	0,56	0,56	0,36	0,46
Mais Rápido	43	0,36	0,56	0,56	0,36	0,46
Mais Rápido	45	0,48	0,56	0,56	0,16	0,44
Mais Rápido	52	0,36	0,56	0,56	0,28	0,44
Mais Rápido	44	0,36	0,56	0,56	0,2	0,42
Mais Rápido	46	0,36	0,56	0,56	0,12	0,4
Mais Rápido	47	0,36	0,56	0,56	0,08	0,39
Ok	64	1	1	1	0,9655	0,991375

Função	Gesto	Peso TOG	Peso TPO	Peso TOV	Peso TOT	PG
Ok	43	0,9655	1	1	1	0,991375
Ok	5	0,8966	1	1	0,931	0,9569
Ok	3	0,931	0,8966	0,8966	0,8966	0,9052
Ok	2	0,8621	0,8621	0,8621	0,8276	0,853475
Ok	11	0,8276	0,7586	0,7586	0,7931	0,784475
Ok	26	0,6897	0,7586	0,7586	0,8621	0,76725
Ok	73	0,8276	0,7586	0,7586	0,7241	0,767225
Ok	21	0,6897	0,7586	0,7586	0,7586	0,741375
Ok	65	0,7586	0,7586	0,7586	0,6552	0,73275
Ok	4	0,7586	0,7586	0,7586	0,6207	0,724125
Ok	70	0,6897	0,8276	0,8276	0,5172	0,715525
Ok	10	0,6897	0,8276	0,8276	0,4483	0,6983
Ok	15	0,5172	0,7586	0,7586	0,6897	0,681025
Ok	62	0,5172	0,7586	0,7586	0,5862	0,65515
Ok	57	0,5172	0,7586	0,7586	0,5517	0,646525
Ok	75	0,6897	0,7586	0,7586	0,3448	0,637925
Ok	29	0,5172	0,7586	0,7586	0,4828	0,6293
Ok	1	0,5172	0,7586	0,7586	0,4138	0,61205
Ok	67	0,5172	0,7586	0,7586	0,3793	0,603425
Ok	32	0,5172	0,7586	0,7586	0,3103	0,586175
Ok	68	0,5172	0,7586	0,7586	0,2759	0,577575
Ok	45	0,5172	0,7586	0,7586	0,2414	0,56895
Ok	66	0,5172	0,7586	0,7586	0,2069	0,560325
Ok	71	0,5172	0,7586	0,7586	0,1724	0,5517
Ok	6	0,5172	0,7586	0,7586	0,1379	0,543075
Ok	74	0,5172	0,7586	0,7586	0,1034	0,53445
Ok	69	0,5172	0,7586	0,7586	0,069	0,52585
Ok	72	0,5172	0,7586	0,7586	0,0345	0,517225
Siga-me	27	1	1	1	1	1
Siga-me	36	0,9444	0,8889	0,8889	0,9444	0,91665
Siga-me	22	0,8333	0,9444	0,9444	0,8889	0,90275
Siga-me	56	0,8889	0,8889	0,8889	0,8333	0,875
Siga-me	50	0,7222	0,7778	0,7778	0,7222	0,75
Siga-me	62	0,6667	0,7222	0,7222	0,7778	0,722225
Siga-me	60	0,6667	0,7222	0,7222	0,5	0,652775
Siga-me	37	0,6667	0,7222	0,7222	0,3889	0,625
Siga-me	31	0,4444	0,7222	0,7222	0,5556	0,6111

Função	Gesto	Peso TOG	Peso TPO	Peso TOV	Peso TOT	PG
Siga-me	59	0,7778	0,4444	0,4444	0,6667	0,583325
Siga-me	58	0,4444	0,7222	0,7222	0,2778	0,54165
Siga-me	54	0,5	0,4444	0,4444	0,6111	0,499975
Siga-me	53	0,4444	0,4444	0,4444	0,4444	0,4444
Siga-me	61	0,2778	0,4444	0,4444	0,3333	0,374975
Siga-me	57	0,2778	0,4444	0,4444	0,2222	0,3472
Siga-me	4	0,2778	0,4444	0,4444	0,1667	0,333325
Siga-me	30	0,2778	0,4444	0,4444	0,1111	0,319425
Siga-me	43	0,2778	0,4444	0,4444	0,0556	0,30555

Apêndice F – Documentos utilizados nos experimentos



Termo de Consentimento Livre e Esclarecido

Título do estudo: Proposta de um vocabulário de gestos robusto e intuitivo para a comunicação Humano-robô.

Pesquisadores responsáveis: Clebeson Canuto dos Santos e Eduardo Oliveira Freire

Instituição/Departamento: Universidade Federal de Sergipe/ Programa de Pós Graduação em Ciência da Computação.

Telefone para contato: (79) 9924-8450

Local da coleta de dados: São Cristóvão

Prezado(a) Senhor(a):

Você está sendo convidado(a) a participar, de maneira totalmente voluntária, de um experimento onde deverá responder um questionário e depois interagir com um robô móvel tipo Pioneer 3-DX.

Antes de concordar em participar deste experimento, é muito importante que você compreenda as informações e instruções contidas neste documento. Os pesquisadores deverão responder todas as suas dúvidas antes que você se decida a participar. Você tem o direito de desistir de participar do experimento a qualquer momento, sem nenhum malefício.

Objetivo do estudo: Estudar os gestos realizados para poder propor um vocabulário de gestos robusto e intuitivo para a comunicação Humano-robô.

Procedimentos: Sua participação neste experimento consistirá na interação a partir de gestos com um robô móvel tipo Pioneer 3-DX. O processo de interação será filmado para posterior análise.

Benefícios: Com esse estudo, poderemos contribuir para a popularização da robótica a partir de interfaces de comunicação mais intuitivas.

Riscos: A realização do experimento representará apenas um risco mínimo, referente à exposição de imagens e dados pessoais, que será minimizado pela garantia de sigilo, uma vez que, divulgaremos apenas dados estatísticos sobre o questionário preenchido. Quanto às imagens, o seu rosto só será visível caso você o permita a partir da resposta dada no item de sigilo que se segue.

Sigilo: As informações fornecidas por você e as imagens terão sua privacidade garantida pelos pesquisadores responsáveis. A exibição dos rostos dos participantes somente ocorrerá mediante manifestação expressa de concordância. Caso contrário, seu rosto será omitido em todas as imagens divulgadas. Para isso, marque a opção que represente sua resposta para a pergunta abaixo:

Você permite que seu rosto apareça nas imagens que poderão ser divulgadas nos resultados desse estudo?

() Sim. Permito o aparecimento do meu rosto nas imagens divulgadas.

() Não. Gostaria de usar máscara durante o experimento.

() Não. Gostaria que meu rosto fosse substituído por uma tarja preta através de um programa de computador.

Ciente e de acordo com o que foi anteriormente exposto, eu _____, estou de acordo em participar deste experimento, assinando este consentimento em duas vias, ficando com a posse de uma delas.

Voluntário

Pesquisador responsável



Questionário

Número: _____

Atividade: _____

Sexo:

- ☐ Masculino
☐ Feminino
☐ N/D

Faixa de Idade:

- ☐ 18-19
☐ 20-29
☐ 30-39
☐ 40-49
☐ 50-59
☐ >=60

Área do Curso:

- ☐ Exatas e da terra
☐ Humanas
☐ Sociais
☐ Saúde
☐ Agrárias
☐ Linguística, letras e artes
☐ Biológicas
☐ Engenharias e Computação

Trabalha?

- ☐ Sim
☐ Não

Qual a frequência de uso de tecnologias?

- ☐ Sempre
☐ Às vezes
☐ Quase nunca

Destro?

- ☐ Sim
☐ Não

Já utilizou gestos para interagir com alguma tecnologia?

- ☐ Sim
☐ Não

Joga vídeo Games?

- ☐ Sempre
☐ Às vezes
☐ Quase nunca

Observações:

Voluntário

Pesquisador Responsável